Draft Comparative Effectiveness Review

Number XX (Provided by AHRQ)

Comparative Effectiveness of Treatment for Depression after Unsatisfactory Response to SSRIs when used as First-line Therapy

Prepared for:
Agency for Healthcare Research and Quality
U.S. Department of Health and Human Services
540 Gaither Road
Rockville, MD 20850
www.ahrq.gov

Contract No. MME223 290-02-0020

Prepared by:
McMaster University Evidence-based Practice Center,
1280 Main St. E., Hamilton, ON

Investigators
Pasqualina (Lina) Santaguida B.Sc.P.T, Ph.D
Glenda MacQueen MD, F.R.C.P.C, Ph.D
Homa Keshavarz B.S.N, Ph.D
Mitch Levine M.D, F.R.C.P.C, FISPE
Joseph Beyene B.Sc, M.Sc, Ph.D
Parminder Raina B.Sc, Ph.D

None of the investigators has any affiliations or financial involvement that conflicts with the material presented in this report.

This report is based on research conducted by the McMaster University Evidence-based Practice Center (EPC) under contract to the Agency for Healthcare Research and Quality (AHRQ), Rockville, MD (Contract No. MME223 290-02-0020). The findings and conclusions in this document are those of the author(s), who are responsible for its contents; the findings and conclusions do not necessarily represent the views of AHRQ. Therefore, no statement in this report should be construed as an official position of AHRQ or of the U.S. Department of Health and Human Services.

The information in this report is intended to help health care decisionmakers—patients and clinicians, health system leaders, and policymakers, among others—make well-informed decisions and thereby improve the quality of health care services. This report is not intended to be a substitute for the application of clinical judgment. Anyone who makes decisions concerning the provision of clinical care should consider this report in the same way as any medical reference and in conjunction with all other pertinent information, i.e., in the context of available resources and circumstances presented by individual patients.

This report may be used, in whole or in part, as the basis for development of clinical practice guidelines and other quality enhancement tools, or as a basis for reimbursement and coverage policies. AHRQ or U.S. Department of Health and Human Services endorsement of such derivative products may not be stated or implied.

Preface

The Agency for Healthcare Research and Quality (AHRQ) conducts the Effective Health Care Program as part of its mission to organize knowledge and make it available to inform decisions about health care. As part of the Medicare Prescription Drug, Improvement, and Modernization Act of 2003, Congress directed AHRQ to conduct and support research on the comparative outcomes, clinical effectiveness, and appropriateness of pharmaceuticals, devices, and health care services to meet the needs of Medicare, Medicaid, and the State Children's Health Insurance Program (SCHIP).

AHRQ has an established network of Evidence-based Practice Centers (EPCs) that produce Evidence Reports/Technology Assessments to assist public- and private-sector organizations in their efforts to improve the quality of health care. The EPCs now lend their expertise to the Effective Health Care Program by conducting Comparative Effectiveness Reviews (CERs) of medications, devices, and other relevant interventions, including strategies for how these items and services can best be organized, managed, and delivered.

Systematic reviews are the building blocks underlying evidence-based practice; they focus attention on the strength and limits of evidence from research studies about the effectiveness and safety of a clinical intervention. In the context of developing recommendations for practice, systematic reviews are useful because they define the strengths and limits of the evidence, clarifying whether assertions about the value of the intervention are based on strong evidence from clinical studies. For more information about systematic reviews, see http://effectivehealthcare.ahrq.gov/reference/purpose.cfm

AHRQ expects that CERs will be helpful to health plans, providers, purchasers, government programs, and the health care system as a whole. In addition, AHRQ is committed to presenting information in different formats so that consumers who make decisions about their own and their family's health can benefit from the evidence.

Transparency and stakeholder input are essential to the Effective Health Care Program. Please visit the Web site (www.effectivehealthcare.ahrq.gov) to see draft research questions and reports or to join an e-mail list to learn about new program products and opportunities for input. Comparative Effectiveness Reviews will be updated regularly.

Acknowledgments

The researchers at the Evidence-based Practice Center would like to acknowledge the following people for their contributions.

We are grateful to our Task Order Officer, Sonia Tyutyulkova for her support and guidance. Members of the technical expert panel were instrumental in the formation of the parameters and goals of this review.

We would also like to thank those who worked so conscientiously, retrieving and screening citations, abstracting data, preparing figures and editing the report: Julianna Beckett, Bryan Cheeseman, Roxanne Cheeseman, Rashmi D'Mello, Henry Ebron, Mary Gauld, Suzanne Johansen, Sara Kaffashian, Dorothy Kendry, Jinhui Ma, Leah Macdonald, Wid Naima, Pam Ren, Rose Revelo, Maureen Rice, Catherine Salmon, Golnoush Taherzadeh, Leila Taherzadeh, Leena Taji.

Technical Expert Panel

Mark Frye, M.D., Ph.D.; Mayo Clinic Department of Psychiatry and Psychology Rochester MN 55905, Consultant and Professor of Psychiatry UCLA Neuropsychiatric Institute, UCLA Medical Center, Los Angeles, CA

Gerald Gartlehner, M.D., M.PH.; Donau-Universität Krems, Department für Evidenzbasierte Medizin und Klinische Epidemiologie, Austria

Bradley N. Gaynes, M.D., M.PH.; Associate Professor of Psychiatry Department of Psychiatry University of North Carolina School of Medicine Chapel Hill, NC

Sidney H Kennedy, F.R.C.P.C., MBBS, M.D.; Clinical Department Head Toronto General Hospital University Health Network, Toronto, ON Canada

Anthony Levitt, Ph.D., MBBS, F.R.C.PC.; Sunnybrook Health Sciences Centre Toronto, ON Canada

Roger S. McIntyre, M.D., F.R.C.P.C.; Head, Mood Disorders Psychopharmacology Unit, University Health Network, Associate Professor of Psychiatry and Pharmacology, University of Toronto, Toronto, ON Canada

David Price, M.D., CCFP, F.C.F.P.; Chair and Associate Professor, Department of Family Medicine, Faculty of Health Sciences McMaster University, Hamilton Health Sciences, Hamilton, ON Canada

Rhonda J Robinson Beale, M.D.; Chief Medical Officer United Behavioral Health Glendale, CA

Madhukar H. Trivedi, M.D.; Department of Psychiatry, University of Texas Southwestern Medical Center at Dallas, Dallas TX

Sandeep Wadhwa, M.D.; State Medicaid Director, Colorado Department of Health Care Policy and Financing Denver CO

John W. Williams Jr., M.D., M.HS.; Professor of Medicine & Psychiatry; Scientific Editor, NC Medical Journal, Duke University and Durham VA Medical Center, Durham, NC

Judy Zerzan, M.D., M.PH.; Colorado Medicaid Medical Director Colorado Department of Health Care Policy and Financing Denver, Colorado 80203

Anthony Levitt, P.hD., M.B.BS., F.R.C.P.C.; Scientist, Sunnybrook Health Sciences Centre, Toronto, ON

EPC Program Director

Parminder Raina, Ph.D.; Professor of Clinical Epidemiology & Biostatistics, McMaster University, Hamilton, ON

AHRQ Contacts

Sonia Tyutyulkova, M.D., Ph.D.; Task Order Officer, Center for Outcomes and Evidence

Contents

Tables

Figures

Comparative Effectiveness of Treatment for Depression after Unsatisfactory Response to SSRIs as First-line Therapy

Executive Summary

The Effective Health Care Program was initiated in 2005 to provide valid evidence about the comparative effectiveness of different medical interventions. The object is to help consumers, health care providers, and others in making informed choices among treatment alternatives. Through its Comparative Effectiveness Reviews, the program supports systematic appraisals of existing scientific evidence regarding treatments for high-priority health conditions. It also promotes and generates new scientific evidence by identifying gaps in existing scientific evidence and supporting new research. The program puts special emphasis on translating findings into a variety of useful formats for different stakeholders, including consumers.

The full report and this summary are available at
www.effectivehealthcare.ahrq.gov/reports/final.cfm

Background

Depression is a complex mental illness that is associated with disability and reduced quality of life for the person with depression, as well as substantial societal burden. Major Depressive Disorder (MDD) is the second leading medical cause of long-term disability, the fourth leading cause of global burden of disease, and predicted to be the second leading cause by 2020.[1,2] Depression exerts a negative impact on physical health; it reduces adherence to medical treatment,[3] reduces participation in preventive activities,[4] and increases the likelihood of risk factors such as obesity,[5] smoking,[6] and sedentary lifestyles.[7] MDD may be associated with immune dysfunction[8-11] and cardiovascular disease,[12-15] endocrine and neurological diseases and a general increase in chronic disease incidence.[16] Mortality rates are high: approximately 4 percent of adults with a mood disorder die by their own hand and about two thirds of suicides are preceded by depression. In adolescents, untreated depression results in significant disability in school performance, interpersonal relationships, risk of suicidal behavior and completion of suicide, risk of early pregnancy, occupational adjustment, and impaired social and family functioning.[17,18]

Pharmacological agents are one of several treatment modalities used for depression, and one of the most frequently utilized classes of antidepressant medications are the selective serotonin reuptake inhibitors (SSRIs). The rate of treatment response following first-line treatment with SSRIs is moderate, however, varying from 40 to 60 percent; remission rates vary from 30 to 45 percent.[19] Up to one third of persons taking antidepressant medications will develop recurrent symptoms of depression while on therapy.[20] The target goal for acute treatment should be

remission, which is defined as a resolution of depressive symptoms. Response to treatment (usually defined as at least a 50 percent reduction in symptom levels[21]) may not be sufficient as a target outcome because residual depressive symptoms are risk factors for relapse and negative predictors of long-term outcome.[22] Remission is defined as a score within the normal range of the symptom scale and is generally considered a preferable goal of treatment.

Clinicians are faced with a number of treatment options following an inadequate response to an SSRI, and these include: 1) changing the dosage or duration of the current SSRI; 2) switching to another SSRI or another antidepressant; 3) switching to a non-pharmacological agent; 4) adding medications (augmenting agents, other SSRI, other anti-depressants); 5) adding a non-pharmacological treatment (such as psychological therapies, exercise, etc.); and 6) combinations of all of these.

Scope and Purpose of this Review

The primary goal of this comparative effectiveness review is to examine the evidence guiding clinical treatment decisions and ultimately to aid clinicians in their care of patients when SSRI therapy for an index episode does not result in an adequate treatment response. The key questions (KQ) are as follows:

Key Question 1.
Among adults and adolescents with major depressive disorder (MDD), dysthymia, and subsyndromal depression, who are started on an SSRI and who are compliant with treatment but fail to improve either fully, partially, or have no response, what is the benefit (efficacy or effectiveness) of monotherapy and combined therapy?
Key Question 1-A. How does the efficacy/effectiveness vary between the different monotherapies and combined therapies?

Key Question 2.
What are the harms of each of the monotherapy or combined therapies among these adults and adolescents? How do the harms compare across different interventions?

Key Question 3.
How do these therapies compare in different populations (e.g., different depressive diagnoses, disease severity, ages, gender, racial and socioeconomic group, and medical or psychiatric co-morbidities)? These subgroups will be considered with respect to the different interventions.

Key Question 4.
What is the range of recommended clinical actions following the failure of one adequate course of SSRI based on current (<5 years) clinical practice guidelines?

Methods

Search Strategy. The search strategy was limited to studies published from 1980 forward to November 2009, as SSRIs first became available for the treatment of depression in the early 1980's. The databases searched were; MEDLINE, Cochrane CENTRAL, PsychINFO, Cochrane Database of Systematic Reviews, EMBASE, CINAHL, and AMED. The grey literature search included systematic searches of relevant citations of websites; health technology assessment agencies (Hayes Inc. Health Technology Assessment), regulatory information (United States Federal Drug Agency (FDA), Health Canada, Authorized Medicines for European Community), clinical trial registries (ClinicalTrials.gov, Current Controlled Clinical Trials, Clinical Study Results, WHO Clinical Trials), grants and federally funded research (including National Institute of Health, Health Services Research Projects in Progress (HSRProj)), abstracts and conference proceedings (Conference Papers Index, Scopus), and the New York Academy of Medicine's Grey Literature Index. Additionally, the sites of specialty organizations were searched for clinical practice guidelines (CPG) and members of the Technical Expert Panel (TEP) were queried for any additional guidelines of relevance. Reference lists of eligible citations and systematic reviews were also searched for potentially relevant citations.

Study Selection. The study populations were eligible if they included adults ($\geq$18 years) or adolescents (12 to 18 years) with major depressive disorder (MDD), dysthymia, or subsyndromal depression, who met the following criteria: 1) currently on SSRI treatment for the index episode at the time of entry into the study, 2) have been judged to have had an "inadequate response" to an SSRI (fluoxetine, citalopram, fluvoxamine, sertraline, escitalopram, and paroxetin) at the time of entry into the study, or 3) the subjects who are recruited for entry into the study to be placed on an SSRI for purposes of monitoring prospectively the adequacy of their response. Studies where subjects who failed to respond to a non-SSRI antidepressant or a non-pharmacological therapy or combination treatment were excluded. Subjects not receiving SSRI at time of entry into the study, and not recruited to evaluate adequacy of response to an SSRI were excluded. Studies where the entire sample included subjects with postpartum depression, bipolar depression, depressive psychosis, dysphoria, mourning syndrome, postoperative depression, premenstrual dysphoric disorder, pseudodementia, puerperal depression, and seasonal affective disorder were excluded. Similarly, studies where the entire sample were subjects with a cerebrovascular accident, dementias (including Alzheimer's disease, vascular dementia, mild cognitive impairment), Parkinson's disease, hypothyroidism, or Cushings' syndrome were also excluded.

Experimental studies and observational studies with comparator groups were included in this review. Study designs with no comparison group (e.g., case series, qualitative studies) were excluded. There were no exclusions based on the types of pharmacological and non-pharmacological interventions, with the exception of electroconvulsive therapy, vagal nerve stimulation, and repetitive transcranial nerve stimulation.

The primary outcomes included remission (free or nearly free of symptoms) and response (partial to complete or up to 50 percent change relative to baseline). Secondary outcomes of interest included speed of response, relapse, quality of life, adherence, return to work, global change as measured by global assessment scales, and external service utilization.

Data Extraction. Relevant fields of information were extracted from individual studies by trained data extractors using standardized forms and a reference guide; a second reviewer verified the accuracy of the data fields reported. Discrepancies were resolved by consensus or consultation. Extracted data included study and population characteristics, eligibility criteria, types of interventions and treatment specifications, and outcomes.

Assessment of Methodological Quality of Individual Studies. We selected the Risk of Bias Tool by the Cochrane Collaboration[23] to assess randomized controlled and controlled clinical trials. Studies were evaluated for adequacy of collecting and reporting harms using the McHarm scale.[24,25] The AGREE II instrument was used to assess the methodological quality of the CPG.[26]

Applicability. Applicability was assessed by considering the key attributes of the population, intervention, comparator and outcome in the context of a wider spectrum of patients in primary care settings; that is in the context of patients that would likely benefit from these interventions in 'real' world conditions. In particular, we considered the characteristics of the included populations, their past mental health history, severity of illness at baseline, the dose of the interventions and comparators, the use of standardized outcomes, treatment duration, and setting.

Rating the Body of Evidence. The overall strength of the body of the evidence was assessed using the modified GRADE approach.[27,28] The strength of the evidence receives one of four grades: high, moderate, low, or insufficient. Grading of the strength of evidence is applied to individual outcomes, which in this review are applied primarily to outcomes of response and remission. The factors considered for impact on the overall grading of the strength of the evidence include: 1) study limitations (predominately risk of bias criteria) and the type of study design (experimental versus observational); 2) consistency of results (degree to which study results for an outcome are similar (variability is easily explained, range of results is narrow); 3) directness of the evidence (assesses whether interventions can be linked directly to the health outcomes); and 4) precision (degree of certainty surrounding an effect estimate for a specific outcome).

Data Synthesis. Qualitative synthesis was undertaken separately for adults and adolescents, and for MDD, dysthymia and subsyndromal depression. Studies were grouped into three categories for treatment strategies that reflected clinical decisionmaking and these included: 1) monotherapy versus monotherapy interventions, 2) monotherapy versus combined therapy treatments, and 3) combined therapy versus combined therapy treatments.
We evaluated the clinical diversity of the study interventions, populations, and outcomes when considering meta-analyzing studies; given the diversity of interventions and populations, summary estimates were not undertaken. Graphs presenting relative risk of individual studies within the various clinical groupings of interventions were prepared to examine differences of effect size.

Results

Description of Eligible Studies and Clinical Practice Guidelines (CPG). From an initial 41,118 citations, 2,765 were screened at full text, and a final set of 39 primary studies (58 publications) and 23 guidelines (29 publications) were eligible for this review. Publications that presented subgroup analyses, secondary analyses, re-analyses, results of different outcomes (not primary outcome measure), or results for different time points on the same study cohort were considered to be secondary records (or companion publications) to the original studies; as such all STAR*D study publications are counted as a single study (with multiple publications).

KQ1. Among adults and adolescents with Major Depressive Disorder, dysthymia, and subsyndromal depression, who are started on an SSRI and who are compliant with treatment but fail to improve either fully, partially, or have no response, what is the benefit (efficacy or effectiveness) of monotherapy and combined therapy?
KQ1 -A: How does the efficacy/effectiveness vary between the different monotherapies and combined therapies?
Thirty-nine unique studies were eligible for key question one. Thirty-seven studies from 57 publications included adults and two studies in 6 publications included adolescents. One study evaluated subjects with subsyndromal depression and another dysthymia; both of these studies showed no differences between groups when comparing monotherapy or combined therapy treatments. The findings for subjects with Major Depressive Disorder (MDD) are summarized below.

Monotherapy vs. monotherapies in Adults. Seven studies in 13 publications compared monotherapy interventions relative to other monotherapy treatments. All participants (n = 1,855) had MDD and were recruited almost exclusively from outpatient settings. The majority of subjects were white, female and middle aged (range 40 and 49 years). The interventions were a minimum of 4 weeks duration and included comparison of two sertraline doses, and switching from 1) citalopram to venlaxafine, buproprion, sertraline or cognitive behavior therapy (CBT); 2) paroxetine to venlaxafine; 3) fluoxetine to olanzapine or mianserin; or 4) from an SSRI to duloxetine (tapering methods). As a group, these seven studies are considered to have moderate risk of bias, with particular problems in randomization and the role of the funding agency. The findings suggest that there is no advantage between different monotherapy treatments (pharmacological or non-pharmacological) for either response to treatment or remission. The exception was a single study that showed lower dose sertraline was superior to the higher dose. There is limited evidence to determine whether a switch to a non-SSRI antidepressant is superior to a switch to another SSRI in patients with inadequate response to an initial SSRI.

Monotherapies vs. combined therapies in Adults. From 31 studies in 57 publications, 29 evaluated the efficacy of monotherapy relative to combined therapies. Participants in the studies (n = 3,989) were all diagnosed with MDD and recruited predominately from outpatient settings. The majority of subjects were middle aged and white (when ethnicity was reported) and females were the majority in almost all studies. Fifteen studies determined failure to the SSRI prospectively and 14 retrospectively. No studies evaluated subjects specifically for failed response to escitalopram or fluvoxatine alone.

All but two studies employed an RCT design and all studies included pharmacological intervention for at least one treatment arm. The majority of studies employed a study design to

have the comparator arm receive ongoing treatment with an SSRI that had been deemed to be not or inadequately effective at a specified point in treatment; far fewer studies employed a design in which patients were switched to a new treatment in at least one study arm.

Four studies had one treatment arm that evaluated a combination therapy that included non-SSRI antidepressants: clomipramine; buproprion; and desipramine. Twenty five from 29 studies evaluated combination therapies that included augmenting agents. From these, only five augmenting agents were evaluated in two or more studies; these included atypical antipsychotics (olanzapine and risperidone), lithium, buspirone, mianserin, and pindolol. Four studies evaluated the use of non-pharmacological interventions including CBT, dialectical behavior therapy, and exercise. Method of randomization, compliance with treatment, and the role of the funder were at high risk of bias for over 75 percent of these studies. Eighteen studies were funded solely by industry, ten by non-industry sources, and one with both. Overall, these studies were rated as having moderate risk of bias.

The majority of studies showed no advantage of the monotherapy comparator treatment relative to combined therapy for the outcomes of response and remission. The exception was with the atypical antipsychotics (olanzapine and risperidone) augmenting agents. There is also limited evidence evaluating the merits of switching to a different antidepressant (monotherapy) relative to adding another treatment (combined therapy). No switched monotherapy was shown to be better than adding another agent.

Combined therapies vs. combined therapies in adults. There were six studies (n = 832) for which there were treatment arms that compared combination therapies to each other. All but one study were RCT. Women were the majority in all studies, and age ranges varied from 37 to 59 years. Only two studies reported racial composition and these subjects were predominately white. Two studies compared different doses of the same combination drug therapies (ziprasidone and lithium). In addition to SSRI, added therapies included lithium, desipramine, buspirone, buproprion, citalopram, clomipramine or CBT. Overall, these studies were rated as having a moderate level of risk of bias, with problems in randomization, reporting compliance, and balancing prognostic indicators. No combination therapy was shown to have any advantage over any other.

Treatment in Adolescents. A single study (Treatment for Resistant Depression in Adolescents (TORDIA)) had extractable data relevant for adolescents 12 to 18 years, who had failed to respond to previous SSRI treatment. The majority of the sample (68 to 72 percent) were females, with an average age of 16 years. Study subjects were randomized to four treatment arms that included venlafaxine alone or combined with CBT, and switch to SSRI (citalopram, fluoxetine, or paroxetine) alone or with CBT. This study was at low risk of bias. There were no differences between the medication groups, but there was a statistically significant difference between the CBT groups in favor of including CBT for all outcomes.

Strength of the Evidence (SOE). The strength of the evidence (SOE) grading was evaluated for the three treatment strategies for adults with MDD who had a previous inadequate response to an SSRI. The evidence for adults with Dysthymia or Subsyndromal depression, or adolescents with MDD had single trials and were not graded.

When considering monotherapy versus monotherapy treatments for persons with MDD, the differing pharmacological and non-pharmacological interventions were considered as a single group, given that so few studies were eligible in this category. Taking into consideration the moderate risk of bias, the imprecision, and the indirectness of the interventions, the quality of evidence was graded as low for both the outcomes of response and remission. Similarly, all interventions within the combined therapies relative to other combined therapies were grouped as one category for grading; the rating of the quality of evidence was also rated as low. The STAR*D trial contributed to both these groupings and impacted the grade given to these therapies for both outcomes. Overall, these SOE grading suggest limited confidence in the effect estimates, and future research will likely change the effect estimates.

The SOE for the studies evaluating monotherapies relative to combined therapies had more eligible studies that were grouped into distinct intervention groups. Since the majority of studies evaluated augmenting agents in the combination treatment arm, we partitioned the studies into relevant subgroups (atypical antipyschotics, buspirone, lithium, mianserin) and also all augmenting agents as a single category. With the exception of atypical antipsychotics (moderate quality) and switching to buspirone **(**low quality), all other groupings for augmenting agents were given a rating of insufficient to permit a conclusion. When considering the grouping of interventions into those where switching to or adding a new SSRI, a non-SSRI and non-pharmacological treatment, the SOE was graded as low; the STAR*D trial contributed to many of the comparisons and impacted the final grade in this category.

KQ2. What are the harms of each of the monotherapy or combined therapies among these adults and adolescents? How do the harms compare across different interventions?
Harms for interventions used in both adults and adolescents with MDD who had failed to respond to SSRI were derived from predominately RCTs that evaluated treatment strategies in this population. No observational studies met the eligibility criteria. A clear trend for harms was difficult to specify across the differing interventions in adults. In general, the majority of harms reported were consistent with those associated with antidepressant use and were likely mild to moderate in nature.

Reporting and collecting of harms was problematic, particularly for pre-defining harms, including serious and severe events, and reporting total number of events per group in studies with adults. The single study evaluating harms in adolescents provided good evidence for harms within this population as the study was at low risk of bias. Severe events and serious events such as suicidality, were inconsistently reported in studies with adult MDD populations. A limited number of studies undertook statistical evaluation comparing harms between groups.

KQ3. How do these therapies compare in different populations (e.g., different depressive diagnoses, disease severity, ages, gender, racial and socioeconomic group, and medical or psychiatric co-morbidities)? These subgroups will be considered with respect to the different interventions?
There are eight studies that undertook stratified or subgroup analyses evaluating factors that may impact treatment outcomes in adults (N = 7), and one for adolescents (N = 1). The effects of baseline severity, previous treatment failure, age, gender, and race were not sufficiently

evaluated and inconsistent in their impact on outcomes in adults. There is some evidence from the STAR*D level 2 cohort that would suggest that persons with concurrent anxiety symptoms have less likelihood of achieving remission. There is some evidence from the TORDIA trial that milder depression, less family conflict, and the absence of suicidal behavior are associated with greater likelihood of a positive treatment response at 12 weeks in adolescents.

KQ4. What is the range of recommended clinical actions following the failure of one adequate course of SSRI based on current (<5 years) clinical practice guidelines (CPGs)?
There were a total of 23 clinical practice guidelines (CPG) sponsored by unique organizations and described in 30 publications. There were seven CPG that were specific only to adolescents, 14 CPG for adults alone, and two CPG applicable to both. Four CPG for adults and three for adolescents did not provide any recommendations for patients with previous inadequate responses. Four guidelines included patients with dysthymia and subsyndromal depression but no recommendations for these subgroups which had failed previous treatment for both adults and adolescents. The majority of CPG did not specify a definition for inadequate response. All CPG were applicable to patients from primary care and outpatient settings.

For adults, the majority of CPG did not specify any type of antidepressant when recommending switching to monotherapy strategies. Increasing the dose and duration was frequently recommended but the interval or change in dose was not specified. When combined therapy was recommended, there was a greater tendency to specify the drug for adding antidepressants. However, there was great variability in the augmenting agents recommended. For adolescents, there was approximately an equal number of CPG that specified which agents to consider for monotherapy and which to consider for combined therapies. Many CPG expressed a preference to commence treatment using non-pharmacological prior to pharmacological treatment. Some adolescent guidelines cited adult evidence as the evidentiary basis for suggesting treatment strategies.

Future Recommendations

Future research should include a broader representation of patients with respect to age and ethnicity. Studies should be more consistent in reporting the manner for determining previous history of failed treatment trials and past episodes of depression and should attempt to determine treatment failure in a prospective manner. Trials of new add-on treatments for patients not responding to an antidepressant medication have not examined whether the add-on agent is equally effective when added to a range of antidepressant classes. There appears to be an assumption among investigators in this field that response and remission will be comparable regardless of the class of background medication and perhaps the clinical or neurobiological data to convincingly support this assumption could be confirmed or re-visited. All standard approaches to treating patients with MDD following inadequate response to an SSRI suffer from a lack of adequate evidence to support clinicians' decisions. Whether the clinician is attempting to optimize the antidepressant medication by changing the dose or duration of the SSRI therapy, switching to a new medication, adding another antidepressant treatment or adding a non-antidepressant agent as augmentation, there is a lack of evidence to guide clinicians and patients in choosing the most appropriate strategy. Much more work is required to determine the most effective ways to optimize short and long-term outcomes for adolescents with depression.

Future clinical trials should conform to CONSORT reporting standards for harms. Severe and serious events (including suicidality) were inconsistently reported and improvement is necessary in this area. Development of future CPGs for adolescents or adults should provide a clear definition of inadequate response for both pharmacological and non-pharmacological treatments and should include standardized methods for establishing this in real world settings. Future CPG recommendations should provide greater clarity with regards to recommended actions and the link with the evidence.

Conclusions

Studies in adults with MDD and an inadequate response to an SSRI included a preponderance of subjects with multiple past depressive episodes and multiple past unsuccessful treatment trials. The generalizability of these data to people with few past episodes of depression may be limited. Studies in adults with MDD and an inadequate response to an SSRI included a high proportion of whites and women and tended to have an average patient age in the early forties. Studies with sufficient sample size to explore whether there are differences in race, gender, or across the age spectrum are needed. The number of studies comparing single medications against each other (monotherapy compared to monotherapy) following an inadequate response to SSRI is extremely limited. Extant studies are limited in type of agents utilized, sample sizes, and population characteristics. There is insufficient evidence to determine whether there is a difference between various single agent therapies for the outcomes of response and remission following inadequate response to an SSRI. There is insufficient evidence to evaluate the benefits of ongoing monotherapy with an SSRI to combination treatment involving the addition of another antidepressant medication to the initial SSRI. There is low grade evidence that comparable results are achieved following switch to an alternate antidepressant medication (monotherapy with a new antidepressant) when compared to adding a non-antidepressant treatment to the initial SSRI (traditional augmentation approach). There is moderate grade evidence that adding an atypical antipsychotic medication to ongoing SSRI treatment is associated with higher response and remission rates compared to adding a placebo to ongoing SSRI treatment following inadequate response to the SSRI. There is insufficient evidence to confirm that there is improvement in response and remission rates following the addition of other augmentation agents. There is insufficient evidence to evaluate the benefits or harms of specific combinations of treatments relative to alternative combinations. There is low grade evidence to support the use of CBT in combination with an antidepressant following inadequate response to an SSRI for adolescents ages 12 to 18, with MDD. There is insufficient evidence to support the use of specific treatments for patients with subsyndromal symptoms following an inadequate response to SSRI medications. There is insufficient evidence to support the use of various treatment approaches for patients with dysthymia who have inadequate response to an SSRI.
A clear trend for harms was difficult to specify across the differing interventions in adults. Harms were well evaluated in the single study in adolescents. Reporting and collecting of harms was problematic, particularly for pre-defining harms including serious and severe events and reporting total number of events per group in studies with adults.

The majority of CPGs in adults were applicable to patients with MDD for outpatient and primary care settings. The majority of CPGs did not specify definitions of "inadequate response". Recommendations for monotherapy, including dose or interval changes, switching to a different SSRI, or a non-SSRI were non-specific as to the drug, interval or dose change. Recommendations for combination therapy tended to endorse switching or adding different classes of antidepressants and augmenting agents. However, there was inconsistency across CPGs with regards to the types of augmenting agents to use.

Chapter 1. Introduction

Background

Depression is a complex mental illness that is associated with disability and reduced quality of life for the person with depression, as well as substantial societal burden. Pharmacological agents are one of several treatment modalities used for depression and one of the most frequently utilized classes of antidepressant medications are the selective serotonin reuptake inhibitors (SSRIs). The rate of treatment response following first-line treatment with SSRIs is moderate, however, varying from 40 to 60 percent; remission rates vary from 30 to 45 percent.[19] Up to one third of persons taking antidepressant medications will develop recurrent symptoms of depression while on therapy.[20] The definition of an adequate response to SSRI medications is not consistently operationalized, but it is generally accepted that a 50 percent decrease in symptom severity constitutes a response.[21] Remission from depression is defined as being free or nearly free of symptoms for the current episode. This review evaluates treatment options for patients who fail to improve fully, partially, or have no response to an SSRI medication.

Epidemiology of Depression in Adults and Adolescents

Major depressive disorder (MDD) is the occurrence of one or more major depressive episodes (MDE). An MDE is defined as a period of at least 2 weeks that is characterized either by depressed mood and/or markedly diminished interest or pleasure in all, or almost all, activities in addition to at least 4 other symptoms.[29] Dysthymic disorder is characterized by a chronically depressed mood and at least two other depressive symptoms, that occur most of the day, more days than not, for at least 2 years. The Diagnostic and Statistical Manual – 4th edition (DSM-IV) includes specifiers that can be used to further describe the characteristics of MDE or dysthymic disorder, such as whether an episode of depression includes psychosis or occurs in the postpartum period. Depression is common in adults and adolescents and is characterized by chronic, recurrent episodes that have significant impact on disability and mortality.

Prevalence of depressive disorders. Kessler reported estimates of 16 percent lifetime prevalence and 7 percent annual prevalence of depression in the United States[30]for adults. These are slightly higher than European prevalence rates of 13 percent lifetime and 4 percent annual.[31] European estimates of the prevalence of dysthymic disorder in adults based on Diagnostic and Statistical Manual – 4th edition (DSM-IV) criteria, are 4 percent lifetime and 1 percent annual.[31] Despite increases in provision of treatment for people with depression[32], a reduction in prevalence has not yet been discernable in those countries where before–after comparisons have been feasible.[33,34] This may be in part because a substantial number of people with depression remain not treated or receive inadequate treatment.[35] The Netherlands Mental Health Survey and Incidence Study assessed episode duration in community residents with new-onset episodes. Although 50 percent of people recovered within three months, the recovery rate flattened over time, and the authors estimated that approximately 20 percent would have episodes lasting longer than 24 months.[36]

The prevalence of MDD in adolescents, 12 to 16 years of age, varies from 4 percent to 8 percent. There is an increased risk of depression following puberty, especially in girls relative to boys (2:1 ratio).[18,37] Prevalence of MDD among adolescents has been reported as high as 20 percent.[18]

Although the literature is limited, the few studies evaluating dysthymia in adolescents report disease prevalence varying from 1.6 percent to 8.0 percent[38] Adolescents with depression have high rates of co-morbid psychiatric conditions (reports vary from 40 to 90 percent), including anxiety and attention deficit hyperactivity disorder and substance abuse problems.[37]

The disease burden associated with MDD, dysthymia. MDD is a leading cause of disability across the world.[2,39] Specifically, depression is the second leading medical cause of long-term disability and the fourth leading cause of global burden of disease, predicted to be second leading cause by 2020. The ongoing transition to a knowledge-based economy is expected to further magnify the impact of MDD on occupational functioning (The Standing Senate Committee on Social Affairs Science and Technology, 2006).[40] Depressive disorders negatively affect quality of life (QOL); 63 percent of respondents with MDD had severe impairment in quality of life (QOL), while 85 percent of those with double depression (MDD and dysthymic disorder) and 56 percent of those with dysthymic disorder had QOL impairment in the severe range.[41] The economic burden of depressive disorders is estimated to be $83.1 billion USD.

The National Comorbidity Survey Replication study in the U.S. found that role impairment in people with MDD was lowest in the occupational domain and highest in the social domain.[30] About 60 percent of respondents with an MDE in the past year reported severe or very severe role impairment. Parental depression has a negative effect on the development of their children and on family dynamics[42,43] and intergenerational effects may amplify the impact of depression on population health.[44]

Depression also has a negative impact on occupational functioning. In one study, depressed workers had significantly greater performance deficits than control workers who had rheumatoid arthritis, with regard to performing mental interpersonal tasks, time management, output tasks, and physical tasks.[45] When depressed workers were compared to workers with rheumatoid arthritis, the depressed employees were almost five times more likely to become unemployed than those with arthritis.[46] Depressed employees are also more likely to become unemployed or miss time at work than physically ill employees.[47]

Depression exerts a negative impact on physical health; it reduces adherence to medical treatment,[3] reduces participation in preventive activities,[4] and increases the likelihood of risk factors such as obesity,[5] smoking,[6] and sedentary lifestyles.[7] MDD may be associated with immune dysfunction,[8-12] cardiovascular disease,[13-15] endocrine and neurological diseases and a general increase in chronic disease incidence.[16] Mortality rates are high; approximately four percent of people with a mood disorder die by their own hand and about two thirds of suicides are preceded by depression.
In adolescents, untreated depression results in significant disability in school performance, interpersonal relationships, risk of suicidal behavior and completion of suicide, risk of early pregnancy, occupational adjustment, and impaired social and family functioning.[17,18]

Clinical Assessment, Management, and Response

Depression is frequently under-diagnosed in primary practice in both adults and adolescents. The World Health Organization (WHO) Psychological Problems in General Health Care study reported that primary care physicians diagnosed only 42 percent of adult patients with major depression. Possible benefits of screening and diagnostic tools to improve detection of depression in primary practice have been examined. Several tools are available for monitoring of depressive symptoms and there were no major differences between these instruments in a comparative study[48]. The Patient Health Questionnaire- 9 item (PHQ-9)[49] or the Quick Inventory of Depressive Symptoms – Self Report (QIDS-SR) appear to be increasing in use, perhaps because of their brevity and strong alignment with DSM-IV. The spectrum of depressive morbidity encountered by primary care physicians is broad. There is also recent evidence suggesting that diagnosis of depression in primary care may not be the major barrier to successful treatment, rather it may be that patients are not receptive to suggested treatment for a condition that was not the reason for the visit to the physician. In primary care, the range of interventions offered may extend from close monitoring of mild episodes without immediate treatment (watchful waiting), through guided self-management,[50] brief psychological or behavioral interventions, pharmacological management and, if needed, referral to more specialized services or hospital admission.

Phases of treatment of major depressive episodes. Based on the work of Kupfer,[51] treatment for MDD is commonly divided into three phases: acute, continuation, and maintenance. Acute treatment is aimed at the elimination of symptoms of depression and restoration of psychosocial functioning. Continuation is a prolongation of treatment from four to nine months, such that the episode of depression is considered completely resolved. For the continuation phase, the treatment aims to return patients to baseline function and quality of life, and to prevent recurrence of symptoms. For the maintenance phase, the treatment goal is to prevent recurrence of new episodes of MDD. In this context, relapse is understood to occur during the continuation phase, but recurrence during the maintenance phase.

The target goal for acute treatment should be remission, which is defined as a resolution of depressive symptoms. Response to treatment, usually defined as at least a 50 percent reduction in symptom levels,[21] may not be sufficient as a target outcome because residual depressive symptoms are risk factors for relapse and negative predictors of long-term outcome.[22] Remission is defined as a score within the normal range of the symptom scale and is generally considered a preferable goal of treatment.

Duration of first line treatment prior to establishing adequate response. Embedded within the decision that a patient has not had an adequate response to treatment, is the issue of defining an adequate duration for that treatment. Antidepressant effect may begin within 1 to 2 weeks of initiation[52-54] and early improvement is a prognostic factor for remission.[55] In STAR*D, 93 percent of patients first achieved response after eight or more weeks, while 41 percent of patients who ultimately remitted first attained remission between four and eight weeks after initiating treatment.[56]

Some guidelines suggest that patients with at least minimal improvement (≥20 percent improvement in scores on a depression rating scale after four to six weeks) should continue with the antidepressant for another two to four weeks before considering additional strategies.[57,58] The American College of Physicians (ACP) recommends that clinicians modify treatment if the patient does not have an adequate response to pharmacotherapy within six to eight weeks of the initiation of therapy for MDD.[59]

Outcomes of Importance. There are a number of outcomes that are used within primary care and psychiatry to assess and monitor response to treatment. These scales include those that are self-report or completed by the clinician. Outcome assessment is usually conducted using validated interviewer-rated scales such as the Hamilton Depression Rating Scale (Ham-D)[60] or the Montgomery Åsberg Depression Rating Scale,[61] although limitations have been recognized to the use of the Ham-D in outpatient populations, it remains widely used.[62] Response is typically defined as >50 percent reduction in scores on these scales, while remission is defined as a score within the normal range.[21,62,63]

Defining inadequate Response. The rate of treatment response following treatment with SSRIs is moderate, varying from 40 to 60 percent[19] and up to two thirds of adult patients will not achieve remission with SSRI treatment[56] Up to one third of adults on drug treatment will develop recurrent symptoms of depression while on therapy.[20] Moreover, there is limited evidence identifying reliable predictors (demographic, clinical, or genetic characteristics) of individual response.[64]

Within their systematic review evaluating the efficacy of treatment for adolescents, Williams et al., (2005)[65] showed that the rates of children failing to respond to an initial trial of SSRIs varied from 31 to 64 percent in eligible studies. Similarly, up to 60 percent of adolescents placed on combined treatment for depression (including pharmacological and behavioral therapies) respond positively to these interventions.[17]

A portion of patients, who have experienced an inadequate response from a clinical perspective, may also go on to be defined as treatment resistant if they also fail to respond to subsequent treatment strategies. Treatment resistance is variably defined but usually refers to patients who have failed at least two trials of medication that have been of adequate dose and duration.[66] Some definitions suggest that the failures should be to medications of different classes, but this is not universally accepted.

Monitoring adherence to antidepressants is sometimes difficult, but non-adherence may account for up to 20 percent of patients classified as having treatment resistant depression.[67] Similarly, there is the potential for pseudo-resistance or non-response to inadequate treatment. All this would suggest the difficulty of defining and capturing subjects who have had treatment failure and related subgroups. It may also reflect heterogeneity across studies evaluating the efficacy of SSRIs within this patient population.

Treatment of inadequate response. Treatment strategies following an inadequate response to an SSRI vary and can include monotherapy or combined therapy. Monotherapy options include: 1) an optimization strategy (increasing the dose or extending the duration of the SSRI); 2)

switching to another SSRI; 3) switching to another class of antidepressants; or 4) switching to a non-pharmacological intervention. Combination or add on therapy options include: 1) combining the SSRI with an augmenting agent; 2) combining antidepressants; 3) combining the SSRI with a non-pharmacological therapy.[68] It is also an option to switch to a new antidepressant and simultaneously combine that antidepressant with a second pharmacological or non-pharmacological treatment. This is sometimes referred to as an acceleration strategy.

Scope and Purpose of this Review

A variety of treatment strategies aimed at helping individuals who have inadequate responses to SSRIs have been studied in patients with depression. The primary goal of this comparative effectiveness review is to examine the evidence guiding clinical treatment decisions and ultimately to aid clinicians in their care of patients when SSRI therapy for an index episode does not result in an adequate treatment response. The key questions (KQ) are as follows:

Key Question 1.
Among adults and adolescents with major depressive disorder (MDD), dysthymia, and subsyndromal depression, who are started on an SSRI and who are compliant with treatment but fail to improve either fully, partially, or have no response, what is the benefit (efficacy or effectiveness) of monotherapy and combined therapy?
Key Question 1-A. How does the efficacy/effectiveness vary between the different monotherapies and combined therapies?

Key Question 2.
What are the harms of each of the monotherapy or combined therapies among these adults and adolescents? How do the harms compare across different interventions?

Key Question 3.
How do these therapies compare in different populations (e.g., different depressive diagnoses, disease severity, ages, gender, racial and socioeconomic group, and medical or psychiatric co-morbidities)? These subgroups will be considered with respect to the different interventions.

Key Question 4.
What is the range of recommended clinical actions following the failure of one adequate course of SSRI based on current (<5 years) clinical practice guidelines?

Chapter 2. Methods

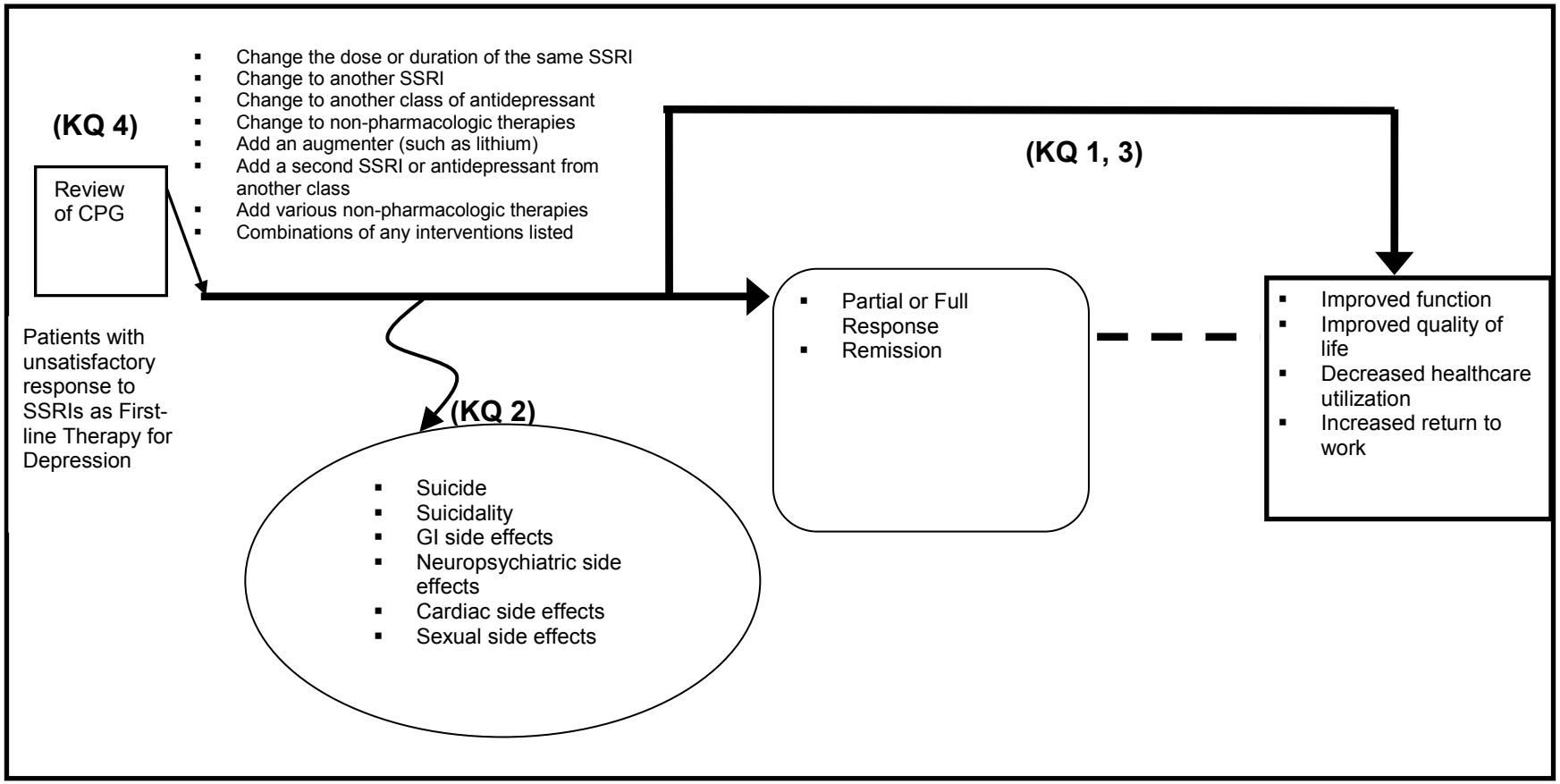
Topic Development

The topic of this report and preliminary key questions were developed through a participatory process involving the public, the Scientific Resource Center (www.effectivehealthcare.ahrq.gov/aboutUS/contract.cfm) for the Effective Health Care program of the Agency for Healthcare Research and Quality (AHRQ), and various stakeholder groups. We communicated with eight key informants who represented psychiatrists, primary care practitioners, consumer representative, and researchers in the area when formulating the research questions. Additional study, patient, intervention, and eligibility criteria, as well as outcomes, were refined and agreed upon through discussions between the McMaster University Evidence-based Practice Center, the Technical Expert Panel (TEP) members, our AHRQ Task Order Officer (TOO), a patient representative, and comments received from the public. Upon completion of the topic development, the key questions were posted for public comment, which were then summarized and discussed with the TEP. Relevant modifications (additions or clarifications) were incorporated.

Analytic Framework

Following consultation with key informants, AHRQ TOO, and our investigative team, we developed our key research questions. Figure 1 shows a flow diagram indicating the relationship between research questions in this Comparative Effectiveness Review (CER). The first box in the figure shows the last question (KQ4) where current (2005 forward) clinical practice guidelines (CPG) are evaluated. The other research questions are related to interventions used following the inadequate response to an SSRI for the index episode of depression. The treatment options following a failed response include the eight options (defined as interventions) for KQ1. Harms associated with any of these interventions are evaluated in KQ2 and can include suicide, sexual dysfunction, gastrointestinal effects, and neuropsychiatric effects. The study effects are evaluated in KQ1, 2 and 3, with the latter question considering subgroups related to different populations with depressive symptoms and other related factors potentially impacting treatment response. We note that intermediate outcomes, such as response and remission may precede quality of life or societal outcomes (costs, utilization).

Figure 1. Analytic Framework



Search Strategy

The search strategy was delimited to studies published from 1980 forward to November 2009, as SSRIs first became available for treatment of depression in the early 1980's. The following electronic bibliographic databases were searched: MEDLINE; Cochrane CENTRAL, PsychINFO, Cochrane Database of Systematic Reviews; EMBASE; CINAHL; AMED. The strategies used combinations of controlled vocabulary (medical subject headings, keywords) and text words. Appendix A details the strategies used to capture relevant citations.

A Grey Literature search was undertaken by the AHRQ Scientific Resource Center (SRC) and identified potentially relevant citations or information by searching the websites as follows: 1) Health Technology Assessment agencies (Hayes Inc. Health Technology Assessment), 2) Regulatory information (United States Federal Drug Agency (FDA), Health Canada, Authorized Medicines for European Community), 4) clinical trial registries (clinical.trials.gov, Current Controlled Clinical Trials, Clinical Study Results, WHO Clinical Trials), grants and federally funded research (National Institute of Health (NIH), HSRPROJ), Abstracts and conference proceedings (Conference Papers Index, Scopus), and the New York Academy of Medicine's Grey Literature Index. Additionally, the sites of specialty organizations for Clinical Practice Guidelines (CPG) were searched and members of the TEP were queried for potentially relevant guidelines.

Review of reference lists of systematic reviews published from 2005 forward was also undertaken. Similarly, the reference lists of eligible studies at full text screening were reviewed for relevant references. Any potentially relevant citations were cross-checked with our citation database and any that were new were retrieved and screened at full text. Our search will be updated (November 2009 to July 2010) following submission of this peer review draft report.

Study Selection

Types of Participants

Subjects who are classified as having failed treatment or having an "inadequate response" are eligible for this review. Treatment failure would ideally be defined as those subjects who are currently on SSRI treatment for the index episode at the time of entry into the study. At that point these subjects have been judged to have had an "inadequate response" at the time of entry into the study or just prior to randomization. An "inadequate response" is established using a standardized instrument, where the scores relative to baseline reflect an improvement of less than 50 percent.[21,69] The term "inadequate response" is therefore synonymous with terms such as "non responders", "failure to respond", and "treatment failure". These terms primarily reflect the perspective of the clinician or researcher. For this CER, the term "unsatisfactory response" was use to reflect the patient's perception of their response to the intervention to treat their depression.

Specific eligibility is as follows: the study populations were eligible if they included adults (≥18 years) or adolescents (12 to 18 years) with Major Depressive Disorder (MDD), Dysthymia, or Subsyndromal Depression, who meet the following criteria:

- Currently on SSRI treatment for the index episode at the time of entry into the study
- Have been judged to have had an "inadequate response" at the time of entry into the study (by any method)
- The SSRIs that patients would not have responded to as a first-line therapy include the following: fluoxetine, citalopram, fluvoxamine, sertraline, escitalopram, and paroxetine

OR

- The subjects who are recruited for entry into the study to be placed on an SSRI for purposes of monitoring prospectively the adequacy of their response; subsequent evaluation includes an intervention for those that have been shown to not respond adequately to the SSRI.

Exclusion

The study populations were not eligible if adults (>18 years) and adolescents (12 – 18 years) with Major Depressive Disorder (MDD), Dysthymia, or Subsyndromal Depression met the following criteria:

- Subjects who are not receiving SSRI at time of entry into the study (including studies that included on antidepressants but not stratified for SSRI subgroup)
- Subjects who are not recruited to evaluate adequacy of response prospectively
- Persons with post-partum depression, bipolar depression, depressive psychosis, dysphoria, mourning syndrome, postoperative depression, premenstrual dysphoric disorder, pseudodementia, puerperal depression, seasonal affective disorder excluded
- Populations for whom the patho-physiological mechanism of depression is not comparable to those diagnosed with MDD including patients having initially sustained a cerebrovascular accident, dementias (including Alzheimer's disease, vascular dementia, Mild Cognitive Impairment), Parkinson's Disease, Hypothyroidism, or Cushings' Syndrome

Types of Interventions

For key questions 1 to 4, the pharmacological and non-pharmacological interventions of interest are as follows:

Selective-Serotonin Reuptake Inhibitors (SSRIs): Fluoxetine (Fluoxetine Hydrochloride, Prozac, Prozac Weekly, Sarafem, Symbyax), Citalopram (Celexa, Citalopram Hydrobromide), Fluvoxamine (Fluvoxamine Maleate, Luvox, Luvox CR), Sertraline (Sertraline Hydrochloride, Zoloft), Paroxetine (Paroxetine Hydrochloride, Paxil, Paxil CR, Pexeva), Escitalopram (Escitalopram, Escitalopram Oxalate, Lexapro)

Non-SSRI Antidepressants: Duloxetine Hydrochloride (Cymbalta), Venlafaxine (Effexor, Effexor XR, Pristiq), Desvenlafaxine Succinate (Pristiq), Phenelzine Sulfate (Nardil), Tranylcypromine Sulfate (Parnate), Emsam (Selegiline), Moclobemide (Manerix), Doxepin (Sinequan, Zonalon, Doxepin Hydrochloride), Clomipramine (Anafranil, Clomipramine Hydrochloride), Amitriptyline (Amitid, Amitril , Elavil, Endep, Etrafon 2-10, Etrafon 2-25, Etrafon-a, Etrafon-Forte, Limbitrol, Limbitrol DS, Perphenazine and Amitriptyline Hydrochloride combinations - Triavil 2-10,Triavil 2-25, Triavil 4-10), Maprotiline (Ludiomil), Desipramine (Norpramin, Pertofrane), Trimipramine (Surmontil, Trimipramine Maleate), Imipramine (Imipramine Hydrochloride, Imipramine Pamoate, Janimine, Pramine, Presamine, Tofranil, Tofranil-pm), Protriptyline Hydrochloride (Vivactil), Agomelatine (Valdoxan), Reboxetine (Edronax, Vestra), Norvale (Mianserin,Bolvidon, Tolvan), Trazodone (Desyrel, Trazodone Hydrochloride, Trialodine), Mirtazapine (Remeron, Remeron Soltab), Nefazodone (Nefazodone Hydrochloride, Serzone), Bupropion (Aplenzin, Bupropion Hydrochloride, Wellbutrin, Wellbutrin SR, Wellbutrin XL, Zyban);

Non-pharmacological and Complementary and Alternative (CAM) Therapies: Cognitive behavioral therapy (CBT), Interpersonal therapy (IPT), and other psychotherapies (Behavior therapy, Interpersonal therapy (IPT), counseling, problem-solving therapy, psychodynamic therapy, bibliotherapy, guided self-help, distraction therapy), Light therapy, Exercise (any type cardiovascular or strengthening or stretching and including yoga, hydrotherapy), , Complementary and Alternative Medicine (CAM) including Whole Body Systems (e.g., Acupuncture), Mind-Body Medicine (e.g., Meditation), Manipulative and Body-Based Practices (e.g., Massage), Energy Medicine (e.g., Reiki); Biologically Based Practices: Dietary supplements and herbal products (e.g., amino acids, vitamins and minerals, Inositol, herbs, methyl-folate [Deplin], omega-3 fatty acids, SAMe).

Augmenters (no formal indication for use as an antidepressant): Buspirone (Buspar), Gepirone (Ariza), Tandospirone (Sediel); Atypical Antipsychotics: Risperidone (Risperdal), Olanzapine (Zyprexa), Quetiapine (Seroquel), Aripiprazole (Abilify), Ziprasidone (Geodon); Psychostimulants: Amphetamine (Adderall), Methylphenidate (Ritalin); Dopamine agonists: Bromocriptine (Parlodel), Cabergoline (Dostinex), Pergolide (Permax), Pramipexole (Mirapex), Ropinirole (Requip), Apomorphine (Apokyn), Rotigotine (Neupro); Other drugs: Lithium, Pindolol, Tryptophan; Anticonvulsants: Carbamazepine (Tegretol), Sodium Valproate, Lamotrigine (Lamictal); Anti-Progestational agents: Mifepristone (Mifeprex); Sex Hormones: Androgens (e.g., Testosterone), Estrogens, Progesterone; Thyroid medications (triiodothyronine,T3), Amisulpride (Solian), Phenytoin (Dilantin, Phenytek), Modafinil (Provigil, Alertec, Modavigil, Modiodal, Modafinil, Carim, Armodafinil, Nuvigil), N-methayl-D aspartate (NMDA) NR2B subunit selective agonist CP-101, 106, mecamylamine hydrochloride (Inversine), Atomoxetine (Strattera).

We excluded studies with electroconvulsive therapy, vagal nerve stimulation, and repetitive transcranial nerve stimulation as the intervention.

For key question 4, we evaluated CPG that focus on guidelines at a national level or from key professional organizations published in English (but not limited to any country).

Types of Comparators

We identified and included studies with comparative intervention groups. From a design hierarchy perspective, comparative group designs provide stronger evidence for efficacy and effectiveness than non-comparative designs.

The interventions (either alone or in combination) may be compared to any of the following:

1. Placebo
2. Same SSRI dose but different MDD population (for example, mild vs. severe MDD)
3. Same SSRI of different dose or duration
4. Other SSRI
5. Other antidepressant (from a different drug class)
6. Non-pharmacological or CAM therapies as described above
7. Adjunct therapy: combination of an augmenter plus SSRI
8. Adjunct therapy: combination of non-pharmacological or CAM therapy plus SSRI
9. Adjunct therapy: combination of augmenter and non-pharmacological or CAM therapy

Outcomes

Primary outcomes include the following:

1) Partial or complete response
   Complete response to treatment is defined as a minimum of 50 percent change relative to baseline using a standardized instrument.[21,69] Partial response refers to a change in baseline score from 25 to 49 percent.
2) Remission
   Remission is defined as being free or nearly free of symptoms. It is typically established by achieving a threshold score using a standardized instrument.

Secondary outcomes include the following:

3) Speed of Response
4) Relapse
   Relapse is defined as a return of symptoms satisfying the full syndrome criteria for an episode and which occurs following a period of remission but before recovery. Relapse is the point at which recurrent symptoms are severe enough that the clinician determines an intervention is warranted. Relapse is related but distinct from the term recurrence. Recurrence is defined as the return of the disease after its apparent cessation (symptoms return after a period of remission).
5) Quality of life
6) Adherence
7) Return to work
8) Global change as measured by global assessment scales
9) External service utilization

Additional Eligibility Criteria

Study Design

Inclusions:

1) Experimental studies with comparator groups (randomized and quasi-randomized trials).
2) Observational studies with comparator groups (retrospective and prospective cohort, case control, and interrupted time series with comparison group).
3) Letters with study data and abstracts

Exclusions:

All other study designs (for example, case series, qualitative studies). Editorials, commentaries, and notes.

Language of Publication

Review of non-English publications were excluded

Contacting Authors for additional data

For studies that included populations that had failed to antidepressants that included SSRI, study authors were contacted via email requesting additional stratified outcome data. Studies, where the authors did not respond or contact could not be established, were excluded.

Timing

There are no restrictions on study eligibility with respect to a minimum treatment interval

Settings

Studies that recruited patients from primary care, outpatient, and inpatient mental health settings were included. There were no exclusions for study setting.

Data Extraction

Relevant fields of information were extracted from individual studies by trained data extractors using standardized forms and a reference guide. Prior to performing the data extraction, a calibration exercise was undertaken using a convenience sample of five included studies. Key study elements were reviewed by a second person (study investigator) with respect to study outcomes, seminal population characteristics (past psychiatric history elements and definition of prior "treatment failure"), and characteristics of the intervention. Disagreements were resolved by consensus.

Extracted data included study characteristics (e.g., first author, country of research origin, study design, sample size, sample size calculation or power estimate); clinical indications; and study duration or length of followup. Details of the patient population included age, gender, racial composition, socio-economic status (income, education), sleeping disturbances or levels, co-morbidities (psychiatric and medical histories, use of complementary and alternative medicines

(CAM) treatments concurrently or historically), definition of treatment failure, severity and duration of the depressive disorder. Details of the study intervention and comparator included the type of intervention/comparator (pharmacological and non-pharmacological and the comparators as listed in the eligibility criteria above), dosage of intervention/comparator (type, dose, method of administration), frequency (and treatment fidelity for psychotherapy related interventions), treatment duration (total duration of care), duration of followup, and characteristics of treatment providers. Characteristics of the outcomes included the type of instrument or scale, primary or secondary outcome status, type of effect measure (endpoint or change score, measure of variance (standard deviation, standard error), etc), definition of "adequate" treatment response, and type of statistical analysis (e.g., intention to treat).

Assessment of Methodological Quality of Individual Studies

We interpret methodological quality to include primarily elements of risk of bias related to the design and conduct of the study. In addition, we evaluated the presence of other key biases, such as the funding bias, and a specific form of selection bias related to "treatment failure" being determined prospectively.

We selected the Risk of Bias Tool by the Cochrane Collaboration[23] to assess randomized controlled trials. The tool contains 12 items that include evaluation of the domains of randomization, blinding, co-intervention, and selective outcome reporting biases. Criteria for evaluation are standardized for these domains. Inconsistency amongst raters was minimized by providing adequate training and standardized instructions; disagreements were resolved by consensus.[70] We had selected the Newcastle Ottawa Quality Assessment Tool[71] to assess risk of bias for observational studies but no study of this design was eligible. Additionally, we evaluated studies for adequacy of collecting and reporting harms using the McHarm Tool;[24,25] this tool has been specifically designed for adverse events and captures domains related to the classification of harms, method of collection (active versus passive), and also the level of withdrawals due to adverse events. We used the AGREE II to assess the methodological quality of the CPG.[26] All tools can be viewed in Appendix B.

A study with low risk of bias was defined as a clinical trial fulfilling six or more of the 12 methodological quality criteria in the Risk of Bias Tool. A study with high risk of bias was defined as fulfilling fewer than six criteria. The classification of individual studies into categories of study limitations (high or low), were used to group studies for grading the strength of the evidence.

Applicability

Applicability was assessed by considering the key attributes of the population, intervention, comparator and outcome in the context of a wider spectrum of patients in primary care settings hat would likely benefit from these interventions in "real" world conditions. In particular, we considered the characteristics of the included populations, their past mental health history, severity of illness at baseline, the dose of the interventions and comparators, the use of standardized outcomes, study duration, and setting.

Rating the Body of Evidence

We assessed the overall strength of the body of the evidence using the modified GRADE approach.[27,28] The strength of the evidence receives one of four grades: high, moderate, low, or insufficient. Grading of the strength of evidence is applied to individual outcomes, which in this review are applied to primarily outcomes of response and remission. A grading of "high" would reflect high confidence that the evidence reflects the true effect and that further research is very unlikely to change our confidence in the estimate of the effect. A grading of "low" would reflect that our confidence is low that the evidence reflects the true effect and the expectation that further research is likely to change with our confidence (in the current literature) or the magnitude of the effect estimate. A grading of "moderate" reflects a moderate level of confidence and additional research may change our confidence. A grading of "insufficient" reflects that the evidence is not available or what evidence is available does not permit a conclusion of substance.

There are several factors that may decrease the overall grading of the strength of the evidence and these include: 1) study limitations (predominately risk of bias criteria) and the type of study design (experimental versus observational), 2) consistency of results (degree to which study results for an outcome are similar (variability is easily explained, range of results is narrow), 3) directness of the evidence (assesses whether interventions can be linked directly to the health outcomes), 4) precision (degree of certainty surrounding an effect estimate for a specific outcome). Additional factors that can be considered when evaluating the strength of evidence can include, dose response, plausible confounding that would decrease the effect, magnitude of the effect, and publication bias. All of these factors were considered when grading the strength of the evidence.

Publication bias

Although our search strategy is comprehensive and includes a grey literature search (including potential sources for unpublished trials), there is always the potential for publication bias. Publication bias is important to assess in reviews with the use of drugs, as there is evidence to suggest that industry sponsorship may lead to negative trials not being published,[72] that reporting of adverse events are more favorable to the funder,[73] and that there may be delay in publication of negative findings.[73] Thus, we will carefully scrutinized studies to determine the presence of selective non-reporting of outcomes (both of benefit and harm).
Our grey literature search was undertaken by the AHRQ Scientific Resource Centre research librarian. Part of this extensive search included a large number of citations to regulatory data bases (such as the Federal Drug Association (FDA)) and clinical trial registries. These sources were searched to identify unpublished or ongoing trials in an attempt to minimize publication bias.

Since we had so few trials on a single intervention (less than 10 studies), we did not undertake production of funnel plots. Moreover, we did not estimate a summary estimate (meta-analysis) of any grouping of interventions and as such no estimates of heterogeneity were undertaken.

Data Synthesis

Qualitative synthesis

For each trial, information on population characteristics (including history of treatment(s) for any previous episodes of depression, age of first diagnosis, etc.), study outcomes (both of benefit and of harm), sample size, settings, funding sources, treatments (type, dose, duration, and provider), methodological limitations, statistical analyses, and any important confounders was summarized in text and summary tables. We stratified the presentation of results based on the type of depressive disorder (MDD, dysthymia, and subsyndromal depression) and by age (adolescents, and adults).

Additionally, we grouped study results: a) according to the index treatment categories (i.e.monotherapy or combined therapies) and the corresponding comparator treatment; b) the specific grouping of the pharmacological (SSRI, non-SSRI, augmenting agents) and nonpharmacological treatment. Forest plots and summary tables were generated to display study primary outcomes of response and remission.

Summary tables were created for CPG stratified by country of origin, where possible.

Quantitative synthesis

The decision to pool individual study results was based on clinical judgment with regards to comparability of study populations, treatments, and outcome measures. Specifically, methodological quality (e.g., high-risk of bias vs. low- risk of bias) and clinical diversity (e.g., characteristics of the study population, gender, disease severity), treatment (pharmacological, non-pharmacological) and intervention duration (2 weeks versus 12 months) and outcome characteristics (e.g., different measuring scales) of individual studies were considered. The extent of heterogeneity was based on the clinical appropriateness of the populations, and interventions.

After the final set of eligible studies were extracted a decision was made to not undertake meta-analyses; rather we displayed the relative risk of individual studies shown on a forest plot for the various clinical groupings of study interventions. We used STATA (Version 10, StataCorp, College Station, Texas, U.S.A.) random effects model to estimate the individual study RR for the outcomes of response and remission.

Subgroup and Sensitivity Analysis

No meta-analyses were undertaken in this CER as study populations, interventons and comparators were not deemed sufficiently similar. However, we considered specific factors in

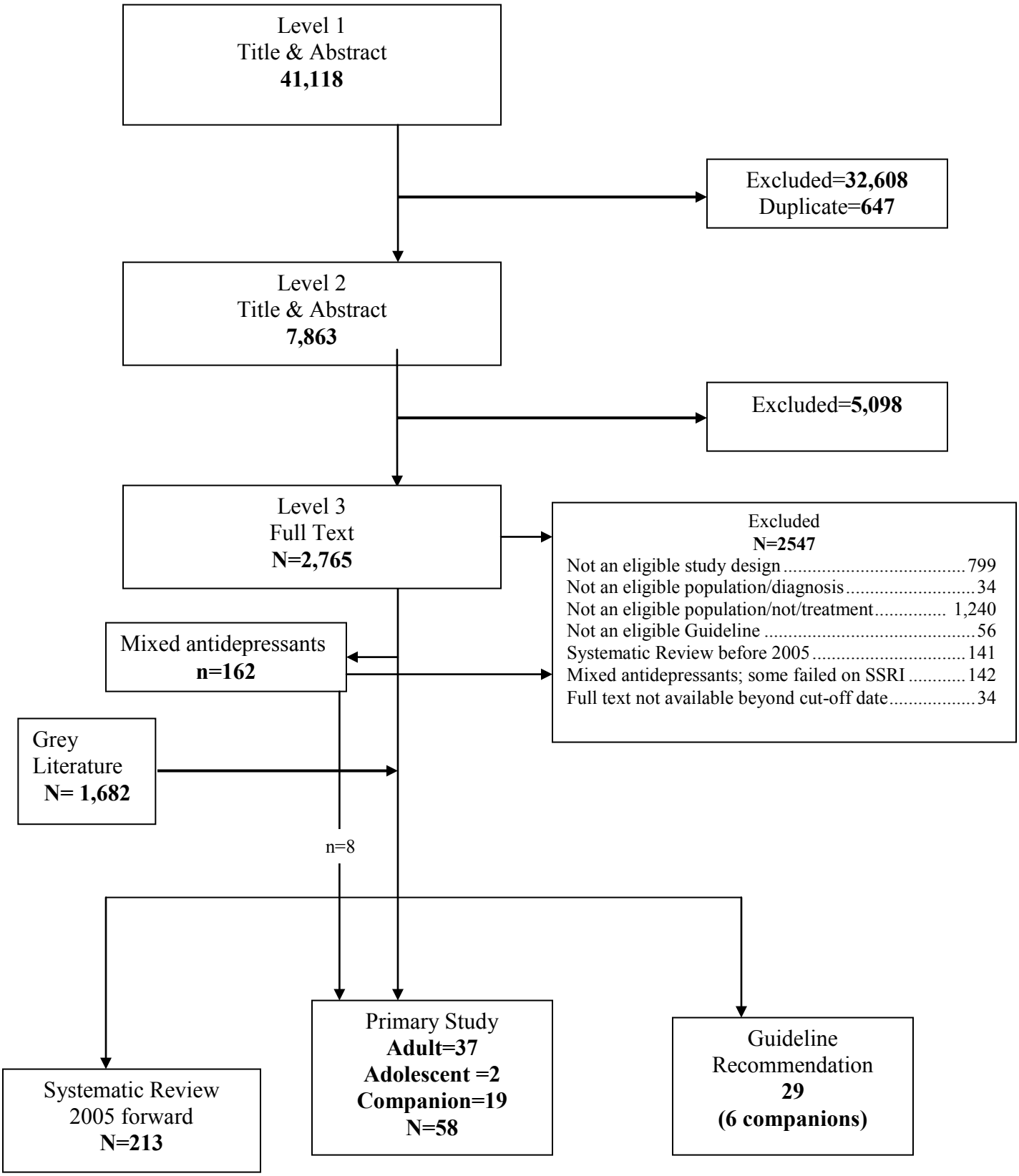
the qualitative presentation of the review findings. Our search yielded only two eligible studies that did not include subjects with MDD, and as such, the impact of the type of depressive disorder could not be explored. Primary studies and guidelines applicable to adults and adolescents were identified and the results were presented stratified by these two age groups. Factors that had the potential to impact study outcomes or account for the clinical heterogeneity, such as gender, number of previous failures, method of determining treatment failure, dose and duration characteristics of the intervention, type of treatment provider were extracted and explored. We summarized these features within the clinical groupings of study interventions monotherapy versus monotherapy, monotherapy versus combined therapy, and combined therapies versus combined therapies. Methodological heterogeneity was also explored within each of these intervention groupings.

Chapter 3. Results

Figure 2 shows the number of citations retrieved in the search from bibliographic and grey literature sources. From an initial 41,118 citations from the seven databases, 647 were duplicates. An additional 1,682 citations were derived from grey literature sources and reviewed for relevancy. Following the initial screen of title and abstract, 32,608 studies were excluded, indicating that the citation was any of the following: 1) a commentary, editorial or narrative review; or 2) not published in English; or 3) not focused on the treatment of depression. At the next level of title and abstract screening, an additional 5,098 citations were excluded as they were: 1) not a primary study, systematic review, or guideline; or 2) not a population with major depressive disorder (MDD), dysthymia, or subsyndromal depression; or 3) evaluated only electroconvulsive therapy, transmagnetic stimulation, and vagal nerve stimulation as treatments for depression. A total of 2,765 citations were then screened at full text. Figure 2 details the reasons for exclusion at full text. Systematic reviews published from 2005 forward were screened for potentially relevant citations that may not have been captured by the search. Thirty nine primary studies (58 publications) were eligible. Twenty-three guidelines in 29 publications were eligible. Reasons for exclusion are detailed within Figure 2.

Publications that presented subgroup analyses, secondary analyses, re-analyses, results of different outcomes (not primary outcome measure), or results for different time points on the same study cohort were considered to be secondary records (or companion publications) to the original studies. For example, there are multiple analyses and publications related to a single study cohort from the Sequenced Treatment Alternatives to Relieve Depression (STAR*D) study. In this study, subjects were evaluated prospectively (level 1) for response to an SSRI, and those who did not respond were followed forward for seven new treatment arms (level 2). Those that failed level two treatments received additional treatments up to level five. STAR*D is included here as a single study with 15 eligible publications; an additional 61 publications based on the STAR*D cohort were excluded because they described results from level 1 only (the prospective evaluation of citalopram efficacy) or they were overview summaries.

Figure 2. Flow of Studies to Final Number of Eligible Studies



Full text screening identified 167 studies with an appropriate design, but for which only a proportion of the sample comprised subjects that were initially treated with an SSRI. In most studies, initial treatment consisted of a variety of possible antidepressants that included, but were not limited to, SSRI medications. The corresponding authors of these studies were contacted by email and asked to provide data stratified for the subgroup treated with an SSRI. Six authors[74-81] provided additional information specific to the SSRI failed subjects and these data are reported in this review. For the remaining 161 studies, 22 authors indicated that they could not provide SSRI failed subject results, 116 did not respond to email contact by the specified cut-off date, and for 23 contact information could not be found Appendix D provides a list of excluded studies and the reasons for their exclusion.

We present the review findings by key question (KQ) and further stratified by adults and adolescents.

Key Question 1. Among adults and adolescents with major depressive disorder (MDD), dysthymia, and subsyndromal depression, who are started on an SSRI and who are compliant with treatment but fail to improve either fully, partially, or have no response, what is the benefit (efficacy or effectiveness) of monotherapy and combined therapy?

Key Question 1-A. How does the efficacy/effectiveness vary between the different monotherapies and combined therapies?

Thirty-nine unique studies were eligible for KQ1. Thirty-seven studies (59 publications) included adults and 2 studies (6 publications) included adolescents. With respect to the studies that evaluated adults, there were 12 publications (7 studies) for which results are not presented. Eight of the 15 STAR*D publications were not extracted; three studies[82-84] presented results for treatment levels three to five, three studies did not present data specific to any treatment,[85-87] and one study[88] presented cost outcomes based on modeling rather than actual cost data. Three studies did not have data extracted[89,90] or partially extracted[91] as they included treatment protocols that evaluated prospective failure to subsequent non-SSRI or combination therapies prior to randomization to a new treatment (similar to level 3 and beyond in the STAR*D cohort). The authors of these studies[90,91] and the authors from three STAR*D publications[82,83,92] were contacted and asked for results specific to the stream of patients that had failed to SSRI prior to the switch to the new intervention being tested; all authors responded and indicated that this data was not available. There were two additional studies[93,94] that were not extracted; these studies used withdrawal designs (maintenance trials), in which subjects who had successfully responded to the combination of an SSRI and an augmenting agent were then randomized to maintain the current treatment or to switch back to monotherapy.

Similarly, from the two eligible studies that evaluated adolescents, only one had data that could be extracted.[17,95,96] A second study[97-99] (3 publications) from the Treatment of Adolescents Study (TADS), indicated that some subjects in the pharmacological arm were evaluated beyond the first prospective failure to an SSRI (phase 2 and 3), but did not present these results. The author was contacted and there was confirmation that this data was not available.

Treatment in Adults who have had inadequate response to an SSRI

KEY MESSAGES
There is low to moderate grade evidence that the addition of an atypical antipsychotic medication is superior to addition of a placebo in patients who have had an inadequate response to an SSRI. There is insufficient evidence that the addition of other augmenting agents is superior to addition of placebo in patients who have had an inadequate response to an SSRI.
There is low quality evidence that response and remission rates following switch to a different antidepressant (monotherapy) are comparable to addition of another treatment (combined therapy) in patients with inadequate response to treatment with an SSRI.
There is insufficient evidence to determine whether a switch to a non-SSRI antidepressant is superior to a switch to another SSRI in patients with inadequate response to an initial SSRI.
Studies to date include a restricted range of patients, with a preponderance of white women between the ages of forty and fifty and a relatively large number of past depressive episodes. Clinicians could benefit from studies that include a higher proportion of younger and older adults, a higher proportion of men, a broader representation of ethnic groups and subjects with a less extensive past illness burden.

There were 37 studies evaluating adults and all but one study included subjects with MDD. Two studies evaluated subjects with subsyndromal depression[100] and dysthymia.[101] As noted previously, five studies[89-91,94,102] and seven STAR*D publications[82,83,85-88,92] did not have data that could be extracted. Additionally, three STAR*D publications[103-105] and two studies[106,107] present results on predictors of failed response in the population of interest and these are presented in key question four. We present the study results for the eligible and extracted studies based on the type of treatment comparisons as follows: 1) monotherapy compared to monotherapy; 2) monotherapy compared to combined therapy; and 3) combined therapy compared to combined therapy. Some studies evaluated more than two treatment arms, and presented monotherapy compared to monotherapy results, as well as monotherapy compared to combined therapy. As such, some studies are included in multiple sections.

Monotherapy treatment compared to monotherapy treatment in MDD

Overview of Study PICOT Characteristics

There were seven studies in 13 publications that compared monotherapy interventions in subjects who had failed to respond to an SSRI. Five studies[91,108-111] had three treatment arms for which two arms compared single interventions directly. The STAR*D study[49,103-105,112-114] (labeled as level 2 subjects within this study), evaluated four monotherapy interventions and one treatment included cognitive behavioral therapy (CBT). One study[115] evaluated two methods of switching to the same anti-depressant. One study[110] compared two doses of a SSRI.

In total, there were 1,855 participants in treatment arms evaluating single interventions within these seven studies. The sample size in these studies varied from 18[109] to 789;[49] the sample sizes per treatment arm varied from 8[109] to 250.[49] Four studies[49,108,110,112,113,115] exceeded a total sample size of 101 and one study[109] had less than 30 subjects.

Population. Women were the majority of subjects in all studies and female gender distributions varied from 60 to 65 percent in three studies[49,103-105,108,110,112-114] to greater than 70 percent.[109,111,115] One study[91] reported gender characteristics for a larger sample (N = 131) but not for the subgroup extracted for this review (N = 41). No studies reported either significant main effects of gender or significant interactions between gender and response rates across treatment groups.

The racial composition varied from 80 percent white [49,112-114] to 100 percent[115] white race subjects. Three studies did not report ethnicity[91,110,111]. There were no differential patterns of response noted to be based on ethnicity.

Mean ages varied from 40 to 44 years in five studies[49,91,108-110] and 45 to 49 years in two studies.[111,115]

Characteristics of the "Inadequate response" for enrolment. Table 1 shows the manner in which failure respond to an SSRI was established in the reported studies. Two studies determined failure retrospectively and study subjects were on an SSRI at the time of entry into the trial.[111,115] Where inadequate response to the SSRI was determined prospectively, fluoxetine,[108,109] citalopram,[116] paroxetine,[91] and sertraline[110] were the SSRIs for which failure was established. No study evaluated subjects specifically for failed response to escitalopram or fluvoxatine alone. Two studies[91,110] excluded subjects with a history of failure over a two week period to any intervention (antidepressant or augmenting agent) used in the current study. One study[115] excluded subjects with a lack of response in the current episode to a serotonin–norepinephrine reuptake inhibitors (SNRI). This study evaluated two methods of switching from an SSRI to duloxetine (an SNRI). No other study evaluating monotherapy treatment excluded or included subjects based on previous failures to any treatment.

Table 1. Method of establishing failure to SSRI and intervention in studies comparing monotherapy strategies following SSRI non-response

	MONOTHERAPY compared to MONOTHERAPY			
	Change dose/ duration of current SSRI	Switch to other SSRI medication	Switch to non-SSRI medication	Switch to non-pharmacological treatment
Prospective trials				
Citalopram		Rush[49]	Rush[49]	Rush[49]
Escitalopram				
Fluoxetine			Thase[108] Shelton[109]	
Fluvoxamine				
Paroxetine			Bondolfi[91]	
Sertraline	Licht[110]			
Retrospective trials				
Medical record/ confirmation clinician				
Patient Self Report				
On specific medication at study entry			Ferreri[111] (Fluoxetine) Perahia[115] (Any SSRI)	

Abbreviations: SSRI= selective serotonin reuptake inhibitors

Mental Health Histories of Study Participants. Three studies, using the Hamilton Depression scale (HAM-D) 17 item version, reported mean baseline scores that varied from 19 (SD 7.3),[49] 21 to 22 (SD 3.3 to 3.9),[115] and 27 to 28 (SD 1.9 to 2.5).[111] One study[110] reported median HAM-D scores of 23 (range 18 to 37). One study[109] reported only that the minimum severity for eligibility was a HAM-D 21 item score of 20 or greater. One study[108] reported Montgomery-Asberg Depression Rating Scale (MADRS) mean score of 30 to 31 (SD 5.9 to 7.0). One study[91] reported baseline scores for a larger sample (N = 131) but not the subgroup of interest (N = 41).

The number of previous depressive episode varied from a median of one episode (range 0 to 8),[91] two episodes (range 0 to 35),[110] or seven to eight episodes (range 12 to 15) in the STAR*D cohort.[49] One study reported that approximately 72 percent of the study subjects had at least one previous episode of depression.[115] Another study[108] reported that 45 percent of the olanzapine group and 79 percent of the fluoxetine group of study subjects had had three or more lifetime episodes. One study[109] did not report the number of previous episodes. A single study reported the proportion of subjects with recurrent depression as 75 percent.[49] Only two studies[91,111] described how previous episodes were defined. Previous episodes were defined as those that required the treatment with antidepressants in both studies. None of the studies specified how information on previous episodes was captured (e.g., by patient report, medical record, etc.).

Length of the current episode was reported as a median value in two studies and a mean in four studies. One study did not report the mean length of the current episode.[109] Median values for length of the current episode varied from eight weeks (range 2 to 52 weeks),[91] and 16 to 20 weeks (range 0 to 960 weeks).[110] Mean values varied from 28 to 32 weeks (range 0 to 42 weeks),[111,115] 52 to 61 weeks (range 78 to 86 weeks)[108] and 118 weeks (SD = 264 weeks).[49] No study specified the manner for collecting length of current episode.

No study in this grouping reported baseline use of complementary and alternative medicines (CAM) at baseline or endpoint.

Intervention and Comparators. All seven studies were labeled as randomized controlled trials (RCT); however the STAR*D cohort had a small proportion of subjects who accepted the randomized arm and as such we classify this as a controlled clinical trial (CCT). The number of treatment arms varied from two[115] to four.[49] Five studies had a prospective run in phase ; the length of this phase varied between 4[91], 6 [109,110] 8[108] and 12 weeks.[49] No study included a washout period before switching to the new interventions. Patient adherence was evaluated in only two studies; one study evaluated this as the number of pills consumed (varied from 94 to 97% adherence),[108] or as not maintaining therapeutic drug monitoring (78 percent adherence).[91]

Table 2 shows the comparison and treatment interventions for the studies evaluating monotherapy. Two studies evaluated switch to sertraline, which represented treatment with a different SSRI; one study[49] used a maximal dose of 200mg/d (titrated from 50mg/d)[49] and one study[110] compared two doses of sertraline (100mg/d and 200mg/d).

Three studies[49,91,115] evaluated a switch to the SNRI venlaxafine, buproprion, and duloxetine. Doses for venlaxafine varied from 37.5 to 375mg/d (extended release)[49] and 150mg/d.[91] Two different methods of switching from the current SSRI to the new medication, duloxetine, were evaluated in one study[115] and as such the dose of 60mg/d was the same for both treatment arms. One study[49] evaluated the use of sustained release buproprion at a maximal dose of 400mg/d (titrated from 150mg/d). This same study had treatment arms for venlaxafine and sertraline.

Two studies[108,109] compared maintenance fluoxetine treatment to olanzapine monotherapy; the doses of fluoxetine were 50mg per day in one study[108] to a range of 20 to 60mg per day in the second study.[109] Olanzapine dosages ranged from 6 to18mg per day in one study[108] and 5 to 20mg/per day in the second study.[109] Another study[111] evaluated mianserin at a dose of 60mg/d.

The comparison arms within these studies included lower doses of sertraline,[110] increased dose of paroxetine,[91] different methods of switching to duloxetine,[115] continuation of fluoxetine at doses for which non-response had been established,[108,109,111] and comparison between sertraline, buproprion, and venlaxafine.[49,113]

Table 2. Monotherapy studies showing the comparison and treatment interventions

Author	SSRI arm	Switch to SSRI
Licht[110] 2002	Sertraline	Sertraline (higher dose)
Switch to non-SSRI		
Perahia[115] 2008	Any SSRI*	Duloxetine (method of taper)
Rush[49] 2006	Venlaxafine	Buproprion
Rush[49] 2006	Sertraline	Venlaxafine
Bondolfi[91] 2006	Paroxetine	Venlaxafine
Switch to Augmenting Agent		
Thase[108] 2007 Shelton[109] 2001	Fluoxetine	Olanzapine
Ferreri[111] 2001	Fluoxetine	Mianserin
Switch to Non-Pharmacological		
Rush[49] 2006 Thase[113] 2007	Venlaxafine/ Sertraline/ Buproprion	CBT

* indicates that comparison arm is not one of the comparison arms after the switch.
Abbreviations: CBT=cognitive behavioral therapy; SSRI= selective serotonin reuptake inhibitors

Primary Outcomes. Two studies indicated that remission was the primary outcome, defined as a MADRS total score of less than 10[91], HAM-D-17 score less than 7 or the Quick Inventory of Depressive Symptoms Self Report (QIDS-SR-16) score less than 5 respectively.[49] A single study indicated that the primary outcome was response based on a 50 percent reduction in the HAM-D-17.[110] Three studies indicated that efficacy (as measured by change score and differences between groups) was the primary outcome, assessed using the HAM-D-17,[115] or MADRS scores.[108,109] One study[111] indicated that both response and remission as determined by the HAM-D-17 were primary outcomes. All studies reported proportions of response (50 percent change relative to baseline) or remission.

Timing of the interventions. Table 3 details the run-in and treatment intervals for the studies comparing monotherapy treatments. The majority of studies evaluated response to the new treatment for six weeks or greater. Similarly, the majority of studies evaluated prospective failure for six weeks or greater.

Table 3. Length of the run-in and treatment phases for all studies

Length of treatment	2/3 weeks	4/5 weeks	6 weeks	8 weeks	>8 weeks
Prospective failure run-in phase		Bondolfi[91]	Licht[110] Shelton[109]	Thase[108]	Trivedi[116] Rush[49] Thase[113]
Prospective failure treatment phase	Bondolfi[91]	Licht[110]	Thase[108]		Trivedi[116] Rush[49] Thase[113] Shelton[109]
Retrospective failure studies			Ferreri[111]		Perahia[115]

Setting. The studies comparing monotherapies were conducted in Europe (Spain, Italy, France, and the United Kingdom),[115] Switzerland,[91] Denmark and Iceland,[110] France,[111] United States and Canada,[108]and the United States (2 studies, 8 publications).[49,103-105,109,112-114]

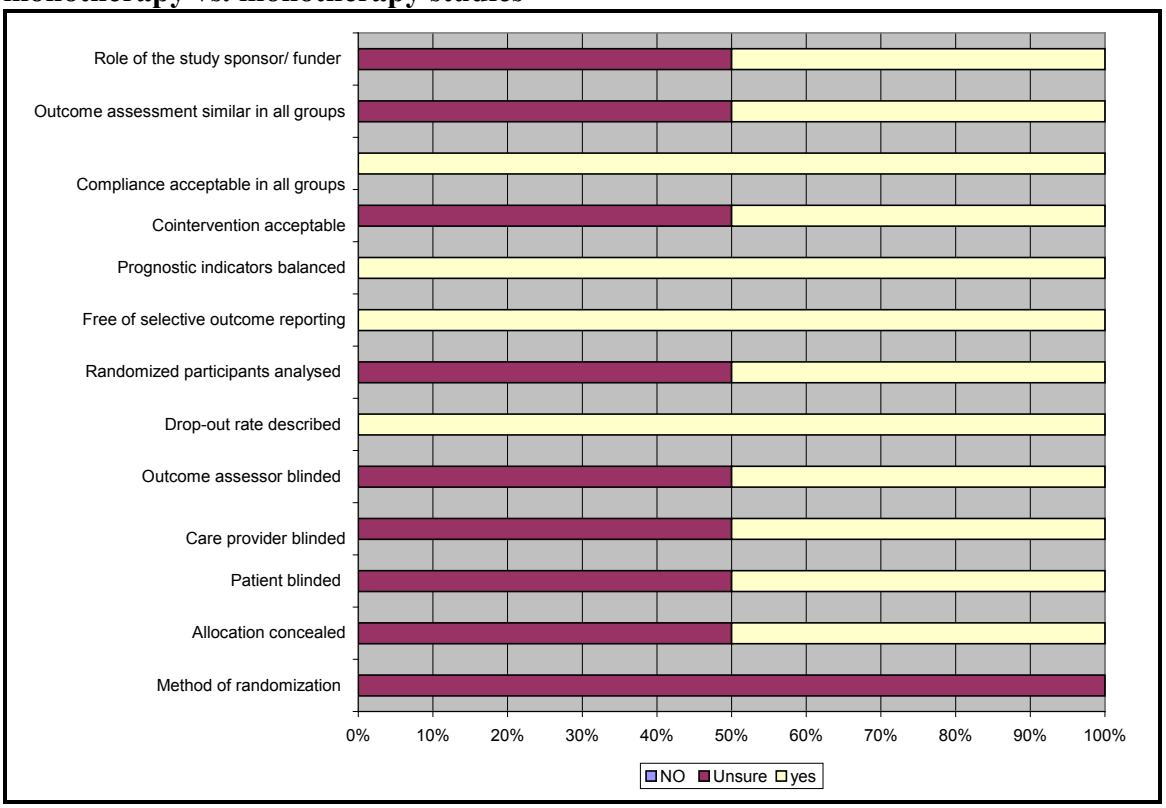
Study participants were all recruited from outpatient psychiatric settings[49,91,108-110,115] and outpatient primary care.[49] A single study recruited from both outpatient and inpatient psychiatric setting.[111]

Risk of Bias

Figure 3.1 shows the distribution of the evaluation for risk of bias using thirteen criteria. None of the seven studies clearly described the method of randomization. All studies were at low risk from biases associated with compliance to treatment, selective outcome reporting, showing reasons for drop outs, and for balancing of important prognostic factors at baseline. For the remaining criteria, only half the studies were at low risk of bias. In particular, the role of the funding agency was not specified in half the studies. All but one study[49] was funded by a pharmaceutical company with a financial interest in one of the products under investigation.

None of the studies included a washout period before randomization to new interventions. Lack of washout in the studies with olanzapine and fluoxetine[108,109] may be problematic as fluoxetine has a long half life (approximately 4 weeks) and the participants are therefore essentially on co-therapy for at least several weeks, even if they are only having olanzapine administered. Most SSRIs have a half life of no more than five days and possibly any very early side effects from the new treatment could actually represent withdrawal from the SSRI if in fact subjects were being switched.As a group these monotherapy studies are considered to have moderate risk of bias given that half of the "risk of bias" items were not met or there was uncertainty.

Figure 3.1. Percent of studies achieving risk of bias using the McHarm criteria in monotherapy vs. monotherapy studies



Efficacy

When single agents were compared against each other following a switch from an SSRI, there was minimal evidence for comparative advantages between treatments. Figure 3.2 reports the rates of response and remission for studies comparing monotherapy treatments. As part of the STAR*D trial,[49,113] few differences were shown for the outcomes of response or remission when patients were switched from citalopram to either another SSRI (sertraline) or a non-SSRI (buproprion or venlafaxine). Similarly, in this same trial, patients who were switched to another monotherapy medication (subgroup) had comparable rates of response to those that were switched to CBT alone. From these studies, only one trial[110] had confidence intervals that did not cross the midpoint, suggesting that the lower dose of 100mg of sertraline plus placebo was superior to 200mg of sertraline plus placebo; response rates of 70 percent compared to 54 percent and remission rates of 38 and 28 percent respectively were reported (Table 4). Almost none of the studies reported any findings from statistical tests to compare the monotherapy arms.

Figure 3.2. Forest plot of monotherapy versus monotherapy interventions for the outcome of response

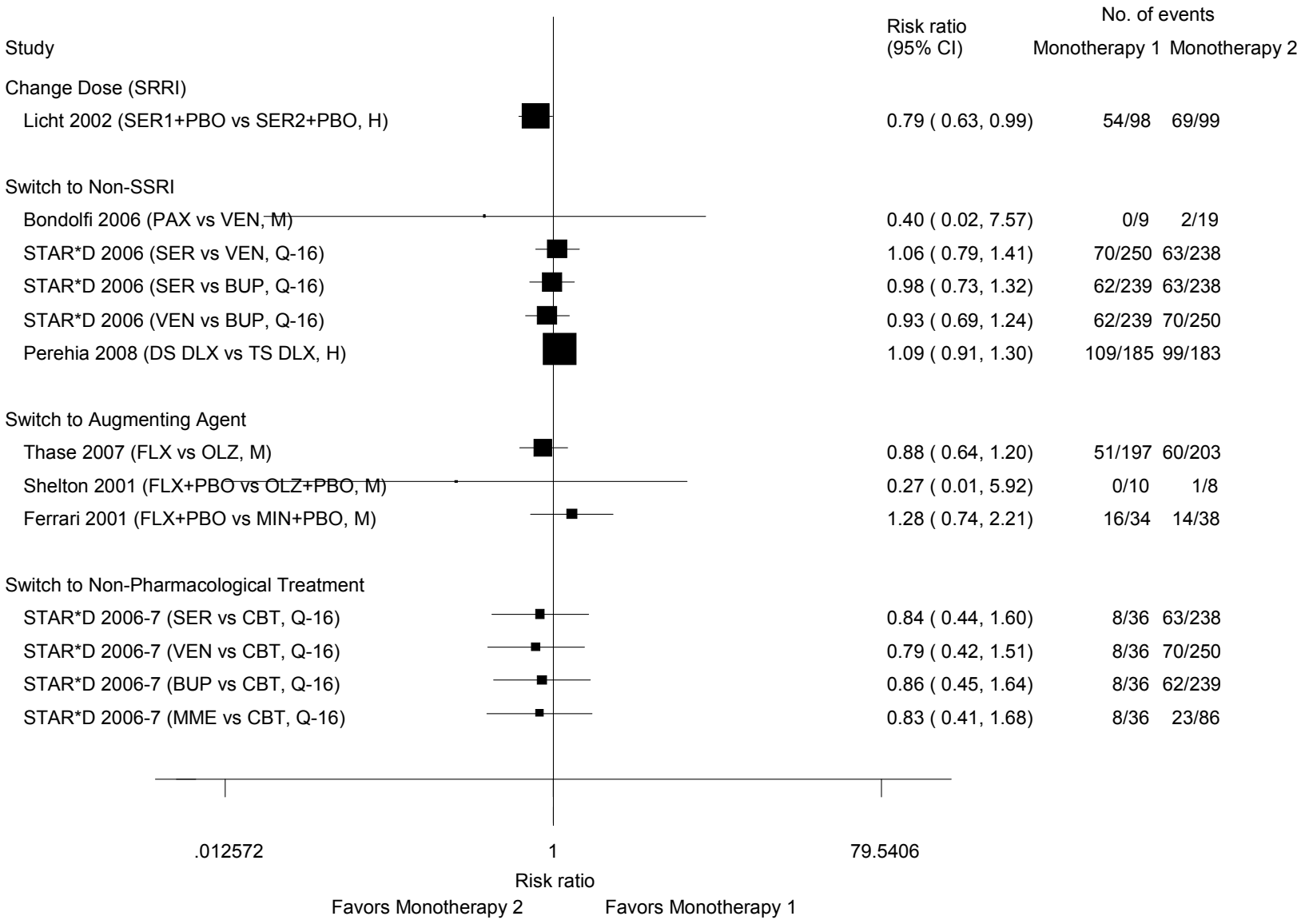


Figure 3.3. Forest plot of monotherapy versus monotherapy interventions for the outcome of remission

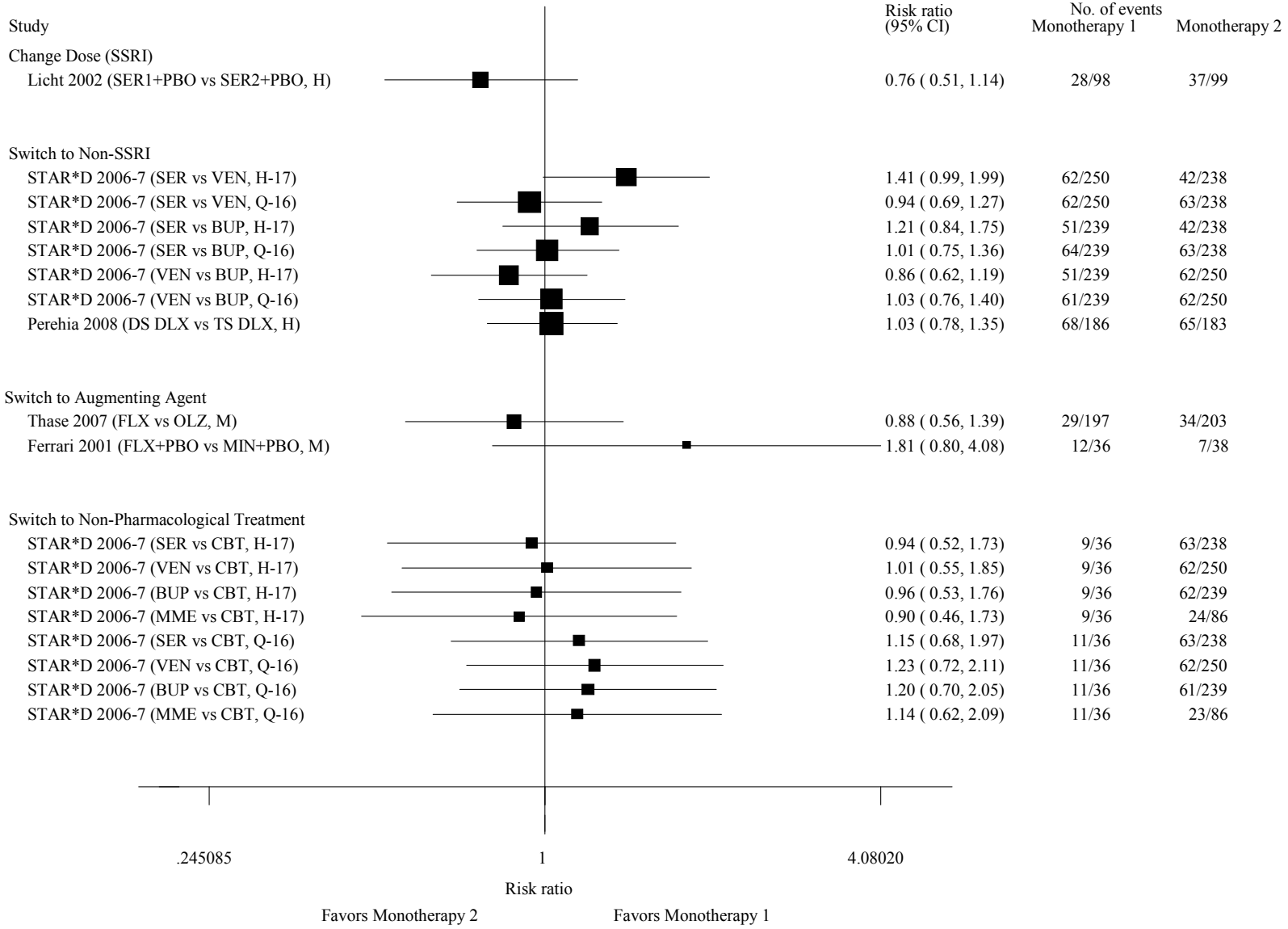


Table 4. Summary of reported rates of response and remission for studies comparing monotherapy treatment to other monotherapy treatments

Study	Duration (weeks)	Rating scale	N*	Comparison and dose (mg/d)	Response N (%)	P value	Remission N (%)	P value
ADD SSRI								
Licht[110] 2002	6	HAMD-NS	99	SER 100 +PBO	69 (70)	0.03	37 (38)	0.19
			98	SER 200 +PBO	54 (64)		28 (29)	
ADD Non SSRI								
Bondolfi[91] 2006	4	MARDS	19	PARO 40mg/d;	2 (10.5)		3 (15.7)	
			9	VEN 150mg/d	0 (0)		0 (0)	
Rush[49] 2006	12	HAMD-17 QUID-SR-16*	238	SER 50-200mg/d	63 (26.7)*		42 (17.6) 63 (26.4)*	
			250	VEN 37.5-375mg/d	70 (28.2)*		62 (24.8) 62 (24.8)*	
			239	BUP 150-400mg/d	62 (26.1)*		51 (21.3) 61 (25.5)*	
Perahia[115] 2008	10	HAMD	183	direct switch DLX 60-120mg/d	99 (54.4)		65 (35.7)	
			185	start-taper switch DLX 60-120mg/d	109 (59.6)		68 (37.2)	
ADD Augmenting Agent								
Thase[108] 2007	8	MADRA	203	FLX 50mg/d;			34 (16.7)	
			197	OLZ 6-18mg/d			29 (14.7)	
Shelton[109] 2001	8	MADRS	8	FLX 20-60mg/d	1 (10.0)	0.11		
			8	OLZ 5-20mg/d	0 (0.0)			
Ferreri[111]	6	HADRS 17	38	FLX 20mg/d	14 (37.0)	0.1	7 (18.4)	0.06

Table 4. Summary of reported rates of response and remission for studies comparing monotherapy treatment to other monotherapy treatments

Study	Duration (weeks)	Rating scale	N*	Comparison and dose (mg/d)	Response N (%)	P value	Remission N (%)	P value
2001			s4	MIN 60mg/d	16 (48.5)		12 (35.2)	
ADD Non-Pharmocological								
Trivedi[116] 2006 Thase[113] 2007	12	HAMD-17 QUID-SR-16*	238	SER 50-200mg/d	63 (26.7)*	NR	42 (17.6) 63 (26.4)*	
			250	VEN 37.5-375mg/d	70 (28.2)*		62 (24.8) 62 (24.8)*	
			239	BUP 150-400mg/d	62 (26.1)*		51 (21.3) 61 (25.5)*	
			36	CBT	8 (22.2)*		9 (25.0) 11 (30.5)*	
			86	Medication	23 (26.7)		24 (27.9) 23 (26.7)*	

Monotherapy versus Combined therapy interventions in MDD

There were 31 unique studies in 57 publications that evaluated monotherapy versus combined therapies. Two studies were withdrawal studies and were not extracted.[94,102] From the remaining 29 studies, four[74,81,113,117] evaluated non-pharmacological interventions combined with SSRI use. The STAR*D study[49,103-105,112-114] for level 2 subjects evaluated four monotherapy interventions and three combined therapies; the CBT monotherapy and citalopram with CBT arms were compared to pharmacological therapies combined and the results are presented in the non-pharmacological section below. A single study[110] evaluated two doses of an SSRI and the same SSRI in combination with an augmenting agent. There were six studies that had subjects who failed to SSRI and non-SSRI antidepressants and subsequently provided some results specific to the failed SSRI group.[74-81] The majority of studies compared the use of a single antidepressant compared to a combined therapy which included an antidepressant with augmenting agents.

In total there were 3,989 participants in studies comparing monotherapy to combined therapies. The total sample size in these studies varied from nine[118] to 1439;[49,103,113,114,116] the sample sizes per treatment arm varied from four subjects[118] to 286.[116] Eleven studies[49,77-80,108,110-114,119-125] exceeded a total sample size of 101 and nine studies[81,109,117,118,126-130] had less than 30 subjects.

Overview of Study PICOT Characteristics

Population. There were two studies evaluating predominately a single gender. One study evaluated males as the intervention was testosterone and it was used as the augmenting agent.[129] Another study evaluated only women being treated with antidepressants alone or combined with exercise.[81] The proportion of women in these studies varied from greater than 70 percent in eleven studies,[49,109-114,118,120-123,126,127,130-133] from 61 to 69 percent in five studies,[108,119,125,134,135] from 51 to 59 percent in two studies[128,136] and from 45 to 49 percent in two studies.[117,124] One study[91] reported gender characteristics for a larger sample (N = 131) but not for the subgroup (step 3A to 3C) extracted for this review (N = 41).

There were six studies for which the authors provided stratified results specific to the subgroup who had failed to adequately respond to an SSRI. However, demographic data was not provided. As such, we have assumed that the SSRI failure subgroup are comparable to the whole sample within the study as they represented over 50 percent of the total sample. When considering the proportion of the study samples who failed to respond adequately to SSRI treatment, there were two studies[75,76] where the sample was 55 to 59 percent, 60 to 69 percent in one study,[81] and greater than 70 percent in three studies.[74,77-79,137] In these six studies, females represented the majority of the subjects in the following proportions; 1) greater than 80 percent in two studies;[74,81] 2) from 70 to 79 percent in two studies;[77,137,138] and 3) from 51 to 60 percent in two studies.[75,76] In the majority of studies, the proportion of men and women per treatment arm were similar with the exception of one small study[128] with 20 subjects which showed differences between groups greater than 10 percent.

Information on racial composition or ethnicity was not reported in 16 studies.[74,81,110,111,118-125,128-132,134,135] For the remaining studies, the majority of subjects were of the white race comprising

between 75 to 89 percent of the sample in six studies[49,75,76,108,112-114,116,117,136] and greater than 90 percent in six studies.[77-79,109,126,127,133,137]

Mean age for the total samples varied from 40 to 44 years in eleven studies,[49,91,108-110,116,123-125,130,135,136] 45 to 49 years in twelve studies,[74-80,111,119-122,127-129,133] 50 to 54 years in two studies,[131,132,134] and greater than 60 years in a single study.[117] One study did not report age characteristics of the very small sample (N = 9).[118] Two studies reported an age range 21 to 54[126] and a range from 40 and 60 years.[81]

Inadequate response. Table 5 shows the manner in which failure to an SSRI had been established. Fifteen studies determined failure prospectively (in an open label manner), and for the majority of these, the subjects were currently on the same antidepressant to which they had shown a poor response. Fourteen studies determined inadequacy of response retrospectively. For studies where inadequate response was determined prospectively, the SSRI to which failure was established included three studies each for fluoxetine,[108,109,124] and sertraline[110,119,136] and two each for citalopram,[116,130] and paroxetine.[91,126] Five studies[75-80,117] used any combinations of SSRI; three studies[75,77-80] specified that fluvoxamine was not one of the SSRI evaluated and these same studies also included escitalopram. No studies evaluated subjects specifically for failed response to escitalopram or fluvoxatine alone.

There were nine studies that excluded subjects because of past failures to specific interventions. Five studies excluded subjects who reported two[123] or three or more previous failures.[77-80,119,126] Three studies[91,110] excluded subjects with a history of failure over a two week period[91] or in the recent episode[110,124] to any intervention (antidepressant, or augmenting agent) used in the current study. Three studies excluded subjects who had an inadequate response to non-pharmacological interventions of electroconvulsive therapy (ECT)[77-80] alone or with repetitive transcranial magnetic stimulation (rTMS), and vagal nerve stimulation (VNS)[131,132] in a previous episode. The remaining twenty studies did not exclude or include subjects based on previous failures to any specific treatment.

Table 5. Method of establishing failure to SSRI in studies comparing monotherapy to combination therapies

	MONOTHERAPY					COMBINED THERAPY			
Determining inadequate response	**Dose or duration change**	**Switch other SSRI**	**Switch non-SSRI**	**Switch to augmenting agent**	**Switch non-pharm**	**Add augmentor**	**Add other SSRI**	**Add non-SSRI AD**	**Add non-pharm**
Prospective									
Citalopram		Rush[49]*	Rush[49]*		Thase[113]*	Trivedi[116]* Baumann[130]		Trivedi[116]*	Thase[113]*
Escitalopram				Thase[108] Shelton[109]					
Fluoxetine						Thase[108] Shelton[109] Fava[124]		Fava[124]	
Fluvoxamine									
Paroxetine			Bondolfi{ 4088			Preskorn[126] Bondolfi[91]			
Sertraline	Licht[110]					Michelson[119] Dunner[136] Licht[110]			
Any SSRI						Mahmoud[76] Keitner[75] Berman[77] Marcus[78] Thase[79,80]			Lynch[117]
Retrospective									
Medical chart									
Self report						George[127]			
Currently on SSRI or other antidepressant			Ferreri[111]			George[127] Shapira[128] Seidman[129] Perry[133] Landén[120-122] Fava[123] Fava[135] Nemets[134] Sokolski[118] Appelberg[125] Ferreri[111]	Altamura[131, 132]	Altamura[1 32] Fava[135]	Carta[81] Wiles[74]

Mental Health History. Table 6 shows the baseline severity reported for the different studies. As expected the baseline scores tended towards the latter quarter of the maximum instrument scores and suggests that subjects had symptoms consistent with those with moderate to severe depression. Two studies did not provide baseline scores.[91,126] The number of previous depressive episode varied from a median of one episode (range 0 to 8),[91] median of two (range 0 to 35);[110] or median of seven to eight (range 12 to 15) in the STAR*D cohort.[49,113,116] Reported mean number of episodes varied from one to two previous episodes[74,133] and three to six.[75,78,80,130,134] Another study[108] reported that 45 percent of the olanzapine group or 79 percent of the fluoxetine group of study subjects had three or more lifetime episodes. Eighteen studies did not report number of previous failed episodes.[76,77,79,81,109,117-129,131,132,135-137]

One study[77,79,80] reported the number of prior adequate antidepressants trials for the current episode and this varied from 67 percent (1 adequate trial) to eight percent (3 adequate trials). Two studies[75,130] showed some differences between treatment groups with respect to previous episodes with the risperidone group having less previous failures. It was not clear in the majority of studies how previous episode were defined and captured (by patient report, medical record, other. No study in this grouping reported baseline use of CAM at baseline or endpoint.

Table 6. Distribution of baseline scores for primary outcomes as a proxy for severity of MDD

	Baseline scores				
Disease specific scale	**10-14**	**15- 19**	**20 - 25**	**26 to 30**	**>31**
MADRS			Appelberg[125]	Thase[108] Landén[121] Landén[120,122] Marcus[78] Goldberg[139] Thase[80] Keitner[75] Appelberg[125]	Dunner[136] Appelberg[125]
BDI				Perry[133]	Wiles[74]
HAM-D -NS		Lynch[117]	Licht[110]	Perry[133]	
HAM-D-31			Fava[123]		
HAM-D-24			Seidman[129]	Nemets[134]	Shapira[128]
HAM-D-21		George[127]	Shelton[109]* Altamura[131] Altamura[132] Sokolski[118] Baumann[130]		
HAM-D-17	Fava[123]	George[127] Keitner[75] Trivedi[116] Rush[49] Thase[113]	Michelson[119] Mahmoud[76] Fava[135] Fava[124] Dunner[136]	Ferreri[111]	
QUIDS-SR16	Trivedi[116] Rush[49] Thase[113]				
Other	Carta[81]				
Total	3	3	12	7	2

Note that two studies[91,126] did not provide baseline scores and some studies provided scores for more than one instrument.

Intervention and Comparator. All but three studies [49,113,116] employed an RCT design with at least some level of blinding. There were four studies[74,81,113,117] that evaluated the use of non-pharmacological interventions including CBT,[74,113] dialectical behavior therapy,[117] and exercise.[81] The remaining 29 used pharmacological agents combined predominately with augmenting agents and a new SSRI or other antidepressants.

Table 7 shows that approximately one quarter of the studies had prospective run in phases and treatment phases that exceeded 8 weeks. Two of the retrospective failure studies provided treatment for this same interval. One study evaluated the Step 3 of the treatment algorithm after only 2 weeks of treatment switch.

Table 7. Details the length of the run-in and treatment phases for all studies

Length of treatment	2/3 weeks	4/5 weeks	6 weeks	8 weeks	>8 weeks
Prospective failure run-in phase		Keitner[75] Bondolfi[91] Baumann[130]	Preskorn[126] Dunner[136] Mahmoud[76] Licht[110] Shelton[109]	Michelson[119] Lynch[117] Thase[108] Fava[124] Berman[77] Marcus[78] Thase [137] Berman[79]	Trivedi[116] Rush[49] Thase[113]
Prospective failure treatment phase	Bondolfi[91]	Preskorn[126] Keitner[75] Licht[110] Baumann[130] Fava[124]	Dunner[136] Thase[108] Mahmoud[76] Berman[77] Marcus[78] Thase [137] Berman[79]	Michelson[119]	Trivedi[116] Rush[49] Thase[113] Shelton[109] Lynch[117]
Retrospective failure studies	Altamura[131] Altamura[132]†	Shapira[128] Wiles[74] Sokolski[118] Landén[120] Landén[121] Landén[122] Nemets[134] Fava[135]	Seidman[129] Perry[133] Appelberg[125] Ferreri[111]	George[127] Carta[81] Fava[123]	

†Indicates treatment was 5 days

Table 8 shows the combined interventions and all other treatment comparisons. Two studies included one treatment arm that evaluated an increased dose of sertraline[110] or the addition of intravenous citalopram.[131,132] Four studies had one treatment arm that evaluated a combination therapy that included non-SSRI antidepressants clomipramine,[131,132] buproprion,[116] and desipramine.[124,135]

The majority of studies evaluated combination therapies that included augmenting agents (25 from 29 studies). From studies with at least one treatment arm using a combination therapy that included an augmenting agent, there were five drugs or classes of drugs for which there was more than one study and these included atypical antipsychotics, lithium, buspirone, mianserin, and pindolol. There were four studies,[91,124,130,135] with at least one treatment arm evaluating the

effect of adding lithium; doses varied from 600mg/d,[124,135] to 800mg/d,[130] and one study did not report the dose.[91] There were four studies evaluating atypical antipsychotics[75,76,108,109] and the doses were similar for studies evaluating olanzapine at 5-6mg/d,[108,109] but varied from 0.5mg/d[75] to 1mg/d in studies assessing risperidone.[76] There were three studies[49,113,116,120-122,125] evaluating buspirone employing final doses that varied from 47mg/d[140] to 60mg/d.[116] Two studies evaluated the use of mianserin[110,111] with doses of 30mg/d[110] and 60mg/d.[111] The augmenting agent pindolol was also evaluated in two studies; the dose was not reported in one study[133] and was 7.5mg/d in the second study.[118]

Table 8. Monotherapy versus combined therapy studies showing the comparison and treatment interventions grouped by type of intervention

Study	Monotherapy	Combined Therapy
Licht,[110] 2002	Sertraline + Placebo	Sertraline (higher dose) + Placebo
Sertraline + Mianserin		
Altamura,[131,132] 2008	SSRI + Placebo (saline)	SSRI + Citalopram (intravenous)
SSRI + Clomipramine (intravenous)		
Add non-SSRI Antidepressant		
Altamura,[131,132] 2008	SSRI + Placebo (saline)	SSRI + Clomipramine (intravenous)
SSRI + Citalopram (intravenous)		
Trivedi,[116] 2006 Rush,[49] 2006* Thase,[113] 2007*	Switching to new monotherapy (Buproprion/Venlaxafine/ Sertraline/ CBT)	Citalopram + Buproprion
Fava,[124] 2002	Fluoxetine + Placebo	Fluoxetine + Desipramine
Fluoxetine + Lithium		
Fava,[135] 1994	Fluoxetine	Fluoxetine + Desipramine
Fluoxetine + Lithium		
Add Augmenting Agent		
Preskorn,[126] 2008	Paroxetine + Placebo	Paroxetine + CP 101,106
George,[127] 2008	Current SSRI +placebo	Current SSRI + Mecamylamine Hydrochloride
Michelson,[119] 2007	Sertraline + Placebo	Sertraline + Atomoxetine
Shapira,[128] 2006	SSRI (fluox/ fluv/ parox) + Placebo	Current SSRI + Phenytoin
Seidman,[129] 2005	SSRI + Placebo injection	Current SSRI + Testosterone injection
Berman,[77] 2007 Marcus,[78] 2008 Thase,[80] Berman,[79]	Switched to new SSRI Escilato/ Fluox/ Sertra/Venlax + Placebo	Switched to new SSRI (Escilato/ Fluox/ Sertra/Venlax) + Aripiprazole
Fava,[123] 2005	SSRI + placebo	SSRI + Modofinil
Nemets,[134] 1999	SSRI + placebo (glucose)	SSRI + Inositial
Dunner,[136] 2007	Sertraline	Sertraline + Ziprasidone 60mg/d
Sertraline + Ziprasidone 80mg/d		
Buspirone		
Appelberg,[125] 2001	SSRI + placebo	SSRI + Buspirone
Landén,[120] 1998 Landén,[121] 2005 Landén,[122] 1999	Citalopram or Paroxetine	Citalopram or Paroxetine + Buspirone
Rush,[49] 2006 Trivedi,[116] 2006 Thase,[113] 2007	Switching to new monotherapy (Sertraline/ Venlaxafine / Buproprion/ CBT)	Citalopram + Buspirone

Table 8. (Cont'd). Monotherapy versus combined therapy studies showing the comparison and treatment interventions grouped by type of intervention

Study	SSRI	Add SSRI
Citalopram + Buproprion/ Citalopram + CBT		
Mianserin		
Licht,[110] 2002	Sertraline Dose 1 + Placebo	Sertraline + Mianserin
Sertraline Dose 2 + Placebo		
Ferreri,[111] 2001	Fluoxetine	Fluoxetine + Mianserin
Lithium		
Baumann,[130] 1996	Citalopram + Placebo	Citalopram + Lithium
Bondolfi,[91] 2006	Paroxetine	Paroxetine + Lithium
Paroxetine *Switch to Venlaxafine*		
Fava,[124] 2002	Fluoxetine + Placebo	Fluoxetine + Lithium
Fluoxetine + Desipramine		
Fava,[135] 1994	Fluoxetine	Fluoxetine + Lithium
Fluoxetine + Desipramine		
Atypical Anti-psychotics		
Thase,[108] 2007	Fluoxetine	Fluoxetine + Olanzapine
Olanzapine		
Shelton,[109] 2001	Fluoxetine + Placebo	Fluoxetine + Olanzapine
Olanzapine + Placebo		
Keitner,[75] 2009	SSRI + Placebo	SSRI + Risperidone
Mahmoud,[76]2007	SSRI + Placebo	SSRI + Risperidone
Pindolol		
Perry,[133] 2004	Fluox/ Setr/ Parox + placebo (lactose powder)	Fluox/ Setr/ Parox + Pindolol
Sokolski,[118] 2004	Paroxetine + Placebo	Paroxetine + Pindolol
Adding Non-pharmacological Treatment		
Wiles,[74]2008	Any SSRI	SSRI + CBT
Carta,[81] 2008	Any SSRI	SSRI + Exercise
Lynch,[117] 2007	Paroxetine/Sertraline/ Fluoxetine	SSRI + Dialectical Behavioral Therapy
Thase,[113] 2007 Rush,[49] 2006 Trivedi,[116] 2006	Switching to new monotherapy (Sertraline/ Venlaxafine/ Buproprion/ CBT)	Citalopram + CBT

*indicates that comparison arm is not the SSRI prior to the switch.

Outcomes. The majority of studies reported change scores as the primary outcome of choice. All but two studies used the MADRS, HAM-D, Beck Depression Index (BDI) or QID-SD-16 for at least one primary outcome; other outcomes used included the clinician global index (CGI),[120-123,125] and the WHOQOL Brief Psychiatric inventory.[81] Only two studies explicitly stated that remission was the primary outcome, defined as a MADRS total score of less than 10,[91] HAM-D-17 score less than 7, or the QIDS-SR-16 score less than 5.[49,113,116] All other studies either specified that the endpoint change score relative to baseline was the primary outcome, or did not report which measure was the primary one to evaluate efficacy.

Setting. The studies were conducted in Denmark and Iceland,[110] Switzerland,[91] France,[111] Italy,[81,131,132] Finland,[125] Norway and Sweden,[120-122] United Kingdom,[74] Israel,[128,134], United States and Canada,[108,136]and United States.[49,75-79,108,109,113,116-119,123,124,126,127,129,130,133,135-137]

Three studies did not report the setting.[108,119-122] From the remaining 26, all studies included subjects from outpatient psychiatric, tertiary or primary care setting with the exception of one study[130] that included patients with a minimum of 4 weeks inpatient hospitalization.

Risk of Bias Assessment

Figure 3.4 shows that method of randomization, compliance with treatment, and the role of the funder were at high risk of bias for over 75 percent of the 29 studies evaluating monotherapies versus combination therapies. Allocation concealment was not achieved by approximately 30 percent of studies. Overall, these studies would be categorized at moderate risk of bias.

Compliance with treatment was evaluated in only three studies[74,91,123] that reported some aspect of compliance with treatment and the remaining studies did not. A single study[125] from 29 employed a washout phase (2 weeks) prior to switching to the new treatment. For studies that employed a switch to a new medication, particularly olanzapine, this may be problematic. For studies switching to other SSRI the half life is approximately five days.

Figure 3.4. Percent of studies achieving risk of bias using the McHarm criteria in studies comparing monotherapy to combined therapy

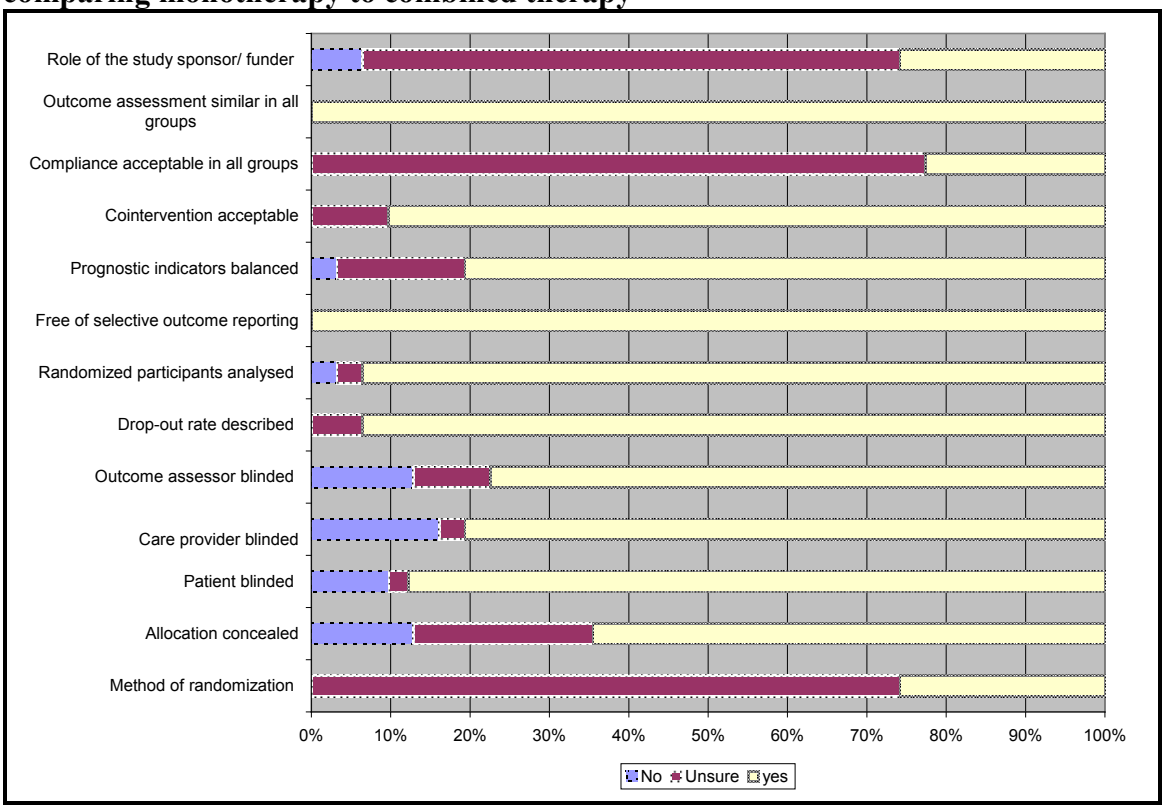


Table 9 shows the distribution of studies for the studies with respect to the source of funding. Eighteen studies were funded solely by industry and ten solely by non-industry sources. One study[127] was funded by both and six studies did not report the source of funding. As indicated in Figure 3.4, the role of the study sponsor was not clearly specified in approximately 75 percent of the 29 studies evaluated here.

Table 9. Sources of funding for studies

	MONOTHERAPY*		COMBINED THERAPY			
Funding source	**Dose or duration change**	**Switch non-SSRI**	**Add augmentor**	**Add other SSRI**	**Add non-SSRI AD**	**Add non-pharm**
Industry		Thase[108] Shelton[109] Licht[110] Ferreri[111] Bondolfi[91]	Preskorn[126] Michelson[119] Landén[121] Landén[120] Landén[122] Berman[77] Marcus[78] Thase[137] Berman[79] Shelton[109] Keitner[75] Mahmoud[76] Fava[123] Appelberg[125] Licht[110] Thase[108] Dunner[136]			
Non-industry	Perahia[115]	Rush[49]	Shapira[128] Shelton[109] Fava[135] Trivedi[116] Fava[124] Baumann[130] Rush[49]		Fava[135] Fava[124]	Wiles[74] Carta[81] Lynch[117] Thase[113]
Both			George[127]			
Not Reported			Seidman[129] Perry[133] Bondolfi[91] Nemets[134] Sokolski[118]	Altamura[131] Altamura[132]	Altamura[131] Altamura[132]	

*No studies reported: Switch other SSRI, Switch to Augmenting Agent, or Switch to Non-Pharm

Efficacy Monotherapy vs. Combined Therapy

Table 10 reports the rates of response and remission reported within the studies in this grouping. Although the majority of studies that could be examined in this review involved comparisons between monotherapy against combination therapy, the wide array of agents used in the combination treatments make identification of trends difficult.

In general, these studies involved one of two study designs. The most commonly employed design involved establishing a cohort of patients who had inadequate response to an SSRI and then randomizing that group to either maintenance of the SSRI and placebo treatment or maintenance of the SSRI in combination with an active intervention. The 'monotherapy' group therefore reflects patients who received ongoing treatment with an SSRI that had been deemed to be ineffective or inadequate at a specified point in treatment. Far fewer studies employed a design in which patients who had an inadequate response to an SSRI were then switched to another treatment and then compared against the combination of the original SSRI plus a new intervention. The STAR*D trial exemplified this latter type of design in which a portion of patients were switched to a new antidepressant treatment following inadequate response to citalopram, while another portion remained on citalopram and had a another treatment added (buspirone, buproprion, or CBT).

In the STAR*D trial, the data did not convincingly support the superiority of either approach (switch to monotherapy or add another treatment, Figure 3.6). Although not statistically significant, there appeared to be a slight, but consistent favoring of the combination treatment approach. In another trial,[131,132] that included a small number of subjects, findings reported that adding either citalopram or clomipramine to another SSRI resulted in greater rates of response compared to patients who received the addition of a placebo; however, the additional treatments were all provided by intravenous infusion over five days. Extrapolation of these results is problematic, as we cannot assume that the addition of the oral compounds would have resulted in a similar pattern of results.

The greatest number of studies in this comparison group involved the treatment strategies of adding an intervention or placebo to ongoing therapy with the SSRI to which patients had shown an inadequate response (Figure 3.5 and 3.8). Note that studies within these figures have been categorized by drug classes (SSRI, non-SSRI, augmenting agents, non-pharmacological). Additionally, we have grouped the studies using augmentation agents based on the number of trials per drug or drug class; interventions that had more than one study included lithium, buspirone, mianserin, atypical antipsychotics, and non-pharmacological therapies. Although there were two studies with pindolol used as the augmenting agent, one did not provide response or remission rates for the monotherapy group.[118]

Overall, none of the augmenting agents showed any advantage over the monotherapy comparator, with the exception of the atypical antipsychotics. Trials of fluoxetine in combination with olanzapine[108,109] and risperidone[76] in combination with SSRI treatment show some relative advantage over monotherapy in patients with MDD for both response and remission. Note that two studies[75,76] provided subgroup data specific to the SSRI failed group and as such the studies were not randomized for this subgroup and balance between groups is therefore not maintained. Two studies[77-80] evaluated the benefits of adding ariprazole in patients who had failed to respond to both SSRI and non-SSRI antidepressants. Although response and remission rates for the SSRI subgroup were not reported, a subgroup analysis indicated that patients on SSRI combined with ariprazole showed consistently greater MADRS total score relative to placebo (-8.6 versus -5.5' treatment difference -3.1 95 percent CI -4.5 to -1.7).

For buspirone and the outcome of remission, we are limited to the different treatment arms of the STAR*D study, and show potentially a small advantage, but this was not for the outcome of response (comparing sertraline relative to citalopram combined with buspirone (see Figure 3.7). This may be an effect of the outcome used to define remission, as no advantage was seen for the QID-SR outcome. Studies evaluating the addition of mianserin show no relative advantage to the monotherapy comparator treatment for either of the outcomes of response or remission.

Evaluation of CBT as an add on therapy showed no advantage when considering any monotherapy comparator; however, most of these data were derived from the STAR*D study
.

Table 10. Summary of reported rates of response and remission for studies comparing monotherapy treatment to combined therapy

Study	Duration (weeks)	Rating Scale	N*	Comparison and Dose (mg/d)	Response N (%)	P value	Remission N (%)	P value
ADD SSRI								
Altamura[131,132] 2008	5 day	HAMD-21	18	SSRI + PBO (saline)	0 (0)	<0.0001		
			18	SSRI + CIT 10mg in 250ml of saline	9 (50)			
ADD NON-SSRI								
Altamura[131,132] 2008	5 day	HAMD-21	18	SSRI + PBO (saline)	0 (0)	<0.0001		
			18	SSRI + CM 25mg in 250ml of saline (intravenous)	11 (61.1)			
Fava[124] 2002	12	HAMD-17	33	FLX 40-60mg/d + PBO	14 (42.4)	0.2		
			34	FLX 20mg/d, DES 25-50mg/d	10 (29.4)			
Fava[135] 1994	4	HAMD-17	15	FLX 40-60mg/d	8 (53)	0.24		
			12	FLX 20mg + DES 25-50mg/d	3 (25)			
Rush[49] 2006	12	HAMD-17 QUID-SR-16*	238	SER 50-200mg/d	63 (26.7)*	NR	42 (17.6) 63 (6.4)*	
			250	VEN 37.5-375mg/d	70 (28.2)*		62 (24.8) 62 (24.8)*	
			239	BUP 150-400mg/d	62 (26.1)*		51 (21.3) 61 (25.5)*	
Trivedi[116] 2006	12	HAMD-17 QUID-SR-16*	279	CIT + BUP, 200-400mg/d	88 (31.8)*		83 (29.7) 108 (38.7)*	0.93 0.16*
Preskorn[126] 2008	6	HAMD-NS	15	PAX 40mg + PBO	3 (20)	<0.10		
				PAX 40mg + CP-101,606 infusion/duration to 1.5 hours and the dose to 0.5mg/kg per hour	12 (80)			
George[127] 2008	8	HAMD-17	10	SSRI + PBO	1 (10)	0.15		
			11	SSRI + ME, 5mg PO/d	5 (45.4)			
ADD AUGMENTING AGENTS								
Michelson[119] 2007	8	MPS	74	SER 100mg/d + PBO			28 (37.8)	0.865
			72	SER 100mg/d + AM 40mg/d			29 (40.3)	
Shapira[128] 2006	4	HAMD-21	9	SSRI + PBO	7 (9)	0.02		
			11	SSRI + PI	2 (11)			
Seidman[129] 2005	6	HAMD-24	13	SSRI + PBO volume matched, (injection)	3 (23.1)	0.226		

Table 10. Summary of reported rates of response and remission for studies comparing monotherapy treatment to combined therapy

Study	Duration (weeks)	Rating Scale	N*	Comparison and Dose (mg/d)	Response N (%)	P value	Remission N (%)	P value
			13	SSRI + TE 200-600mg/d	7 (53.8)			
Berman[77] 2007	8	MADRS	176	New SSRI + PBO	NR for SSRI subgrup		NR for SSRI subgrup	
			182	New SSRI + ARI, 5-15	NR for SSRI subgroup		NR for SSRI subgroup	
Marcus[78] 2008	8	MADRS	191	SSRI + PBO	NR for SSRI subgrup		NR for SSRI subgrup	
			190	SSRI + ARI 5-20mg/d (5-11mg ?)	NR for SSRI subgroup		NR for SSRI subgroup	
Fava[123] 2005	8	HAMD- 17 CGI-1*	153	SSRI + PLB	48 (32)*	>0.09	55 (36)	0.2
			158	SSRI + MOD 100-200	64 (41)*		68 (44)	
Nemets[134] 1999	4	HAM-D 24	18	SSRI original dose + PBO	NR			
			18	IN 12gm/d, SSRI original dose;	NR			
Dunner[136] 2007	8	MADRS	20	SER 100-200mg/d	4 (19)	NS		
			21	SER 100-200mg/d + ZI 40-80mg/d	6 (32)			
			19	SER 100-200mg/d + ZI 80-160mg/d	2 (10)			
Adding Atypical Antipsychotics								
Thase[108] 2007	8	MADRS	203	FLX 50mg/d			34 (16.7)	
			197	OLZ 6-18mg/d			29 (14.7)	0.012
			198	OLAN 6-18mg/d + FLX 50mg/d			24 (27.3)	0.003
Shelton[109] 2001	8	MADRS	8	FLX 20-60mg/d	1 (10)	0.006 vs. com		
			8	OLZ 5-20mg/d	0 (0)	0.003 vs. com		
			10	OLZ 5-20mg/d, FLX 20-60mg/d	6 (60)	0.007		
Mahmoud[76] 2007	6	HAMD-17	74	SSRI + PBO	18 (24.3)	NR	4 (5)	NR
			82	SSRI + RIS 0.25-1mg/d	37 (45.70		21 (25)	
Keitner[75] 2009	4	MADRS	22	SSRI dose maintained + PBO			5 (22.7)	0.011
			47	RIS 0.5-3mg/d + antidepressant dose maintained			24 (51)	
Adding BUS								
Rush[49] 2006	12	HAMD-17 QUID-SR-16*	238	SER 50-200mg/d	63 (26.7)*	NR	42 (17.6) 63 (6.4)*	
			250	VEN 37.5-375mg/d	70 (28.2)*		62 (24.8) 62 (24.8)*	

Table 10. Summary of reported rates of response and remission for studies comparing monotherapy treatment to combined therapy

Study	Duration (weeks)	Rating Scale	N*	Comparison and Dose (mg/d)	Response N (%)	P value	Remission N (%)	P value
			239	BUP 150-400mg/d	62 (26.1)*		51 (21.3) 61 (25.5)*	
Trivedi[116] 2006	12	HAMD-17 QUID-SR-16*	286	CIT + BUS, 200-400mg/d	77 (27)*		86 (30.1) 94 (32.9)*	0.93 0.16*
Appelberg[125] 2001	6	MADRS	51	CIT 40mg/d/FLX 35.4mg/d + PBO	16 (31)	0.034		
			51	CIT 40mg/d/FLX 35.4mg/d + BUSP 35-47mg/d	17 (33)			
Landén[120] 1998	4	CGI-S, CGI-I	60	CIT 46.1mg/d or PARO 39.8mg/d + PBO	28 (46.7)	NS		
			57	CIT 46.1mg/d or PARO 39.8mg/d + BUSP 49mg/d	29 (50.9)			
Adding Li								
Fava[135] 1994	4	HAMD-17	15	FLX 40-60mg/d	8 (53)	0.24		
			14	FLX 20mg/d + LI 300-600mg/d	4 (29)			
Fava[124] 2002	12	HAMD-17	33	FLX 40-60mg/d + PBO	14 (42.4)	0.2		
			34	FLX 20mg/d, LI 300-600mg/d	8 (23.5)			
Baumann[130] 1996	4	HAMD-21	14	CIT 40-60mg/d	2 (14)	0.05		
			10	CIT 40-60mg/d, LI 800mg/d;	6 (60)			
Bondolfi[91] 2006	4	MADRS	19	PARO 40mg/d	2 (10.5)	NR	3 (15.7)	NR
			9	VEN 150mg/d	0 (0)		0 (0)	
			13	PARO 30mg/d + LI	1 (7.8)		0 (0)	
Adding PIDOLOL								
Perry[133] 2004	6	HAMD	17	SSRI + PBO FLX 20-60mg, PARO20-40mg, SER 50mg	6 (35.2)	1		
			21	SSRI + PI Total = only SSRI doses given, PI dose not reported; Group 1 = FLX 20-60mg, PARO 20mg, SER 150-200mg	5 (23.8)			
Sokolski[118] 2004	4	HAMD	5	PARO 40mg/d + PBO	NR	0.001		
			4	PARO 40mg/d + PI 7.5mg/d	3 (75)			
Adding MIN								
Licht[110] 2002	6	HAMD-NS	99	SER 100mg/d + PBO	69 (70)	0.64	37 (38)	0.38
			98	SER 200mg/d + PBO	54 (64)	0.03	28 (29)	0.19
			98	SER 100mg/d + MIN	66 (67)		43 (44)	

Table 10. Summary of reported rates of response and remission for studies comparing monotherapy treatment to combined therapy

Study	Duration (weeks)	Rating Scale	N*	Comparison and Dose (mg/d)	Response N (%)	P value	Remission N (%)	P value
Ferreri[111] 2001	6	HAMD 17	38	FLX 20mg/d	14 (37)	0.1	14 (36)	0.06
			34	MIN 60mg/d	16 (48.5)		6 (18)	
			32	FLX 20mg/d + MIN 60-60mg/d	20 (62.5)		14 (44)	
Adding Non-Pharmacological								
Carta[81] 2008	32	WHOQOL-Bref Psych;	10	SSRI	NR		NR	
			20	SSRI + Exercise	NR		NR	
Lynch[117] 2007	54	HAMD-NS	12	SSRI			6 (50)	NR
			20	SSRI + Dialectical Behavioural Therapy			12 (60)	
Wiles[74] 2008	16	BDI	9	SSRI	0			
			14	SSRI + CBT	8 (56)			
Rush[49] 2006	12	HAMD-17 QUID-SR-16*	238	SER 50-200mg/d	63 (26.7)*	NR	42 (17.6) 63 (6.4)*	
			250	VEN 37.5-375mg/d	70 (28.2)*		62 (24.8) 62 (24.8)*	
			239	BUP 150-400mg/d	62 (26.1)*		51 (21.3) 61 (25.5)*	
			36	Medications Monotherapy	8 (22.2)		9 (25) 11 (30.5)	
			86	CBT	23 (26.7)		24 (27.9) 24 (27.9)	
			65	CIT + CBT	23 (35.4)		15 (23) 20 (20.7)	

Figure 3.5. Forest plot showing monotherapy versus combined therapy for the outcome of response for augmenting agents

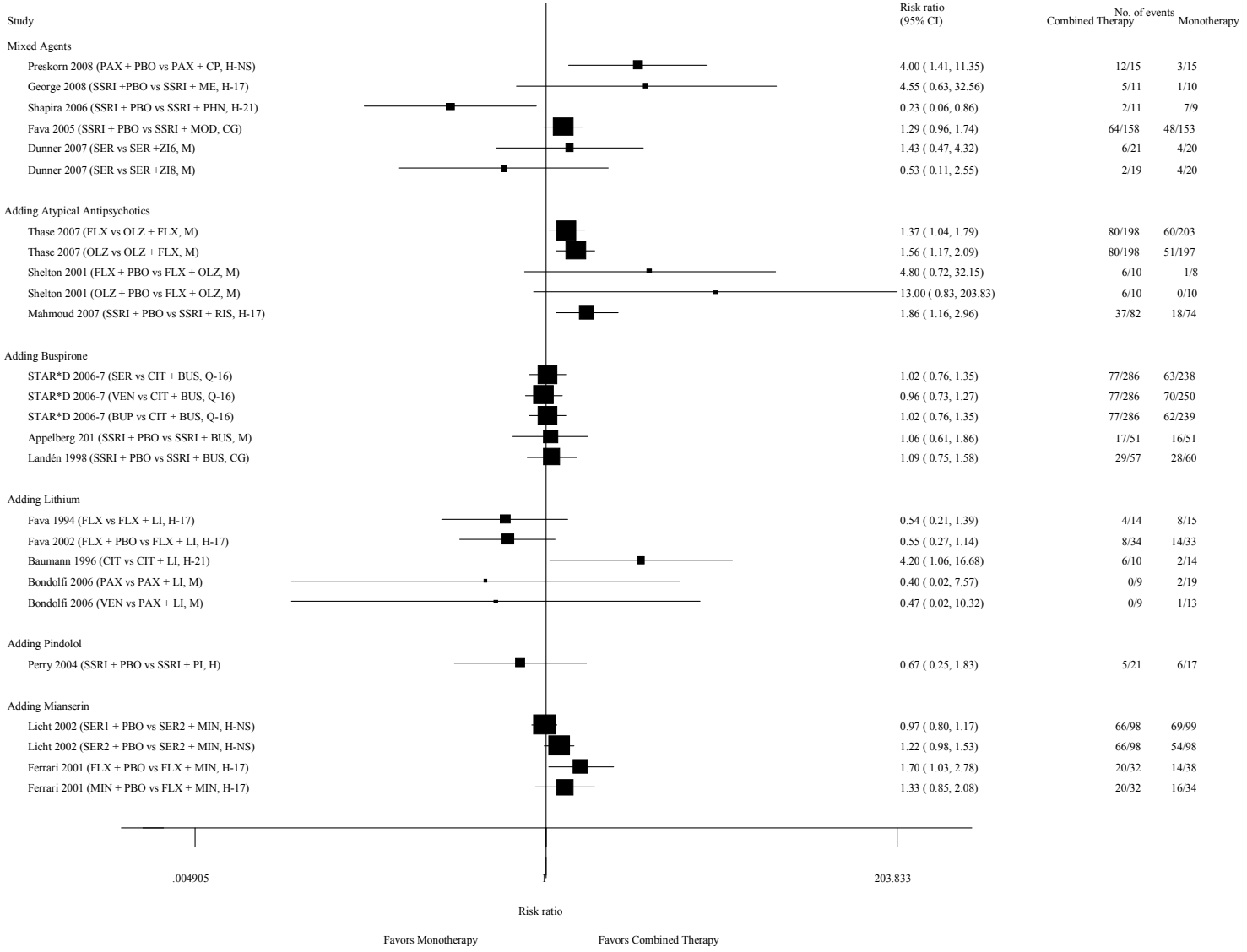


Figure 3.6. Forest plot of monotherapy versus combined therapies for the outcome of response for all interventions but augmenting agents

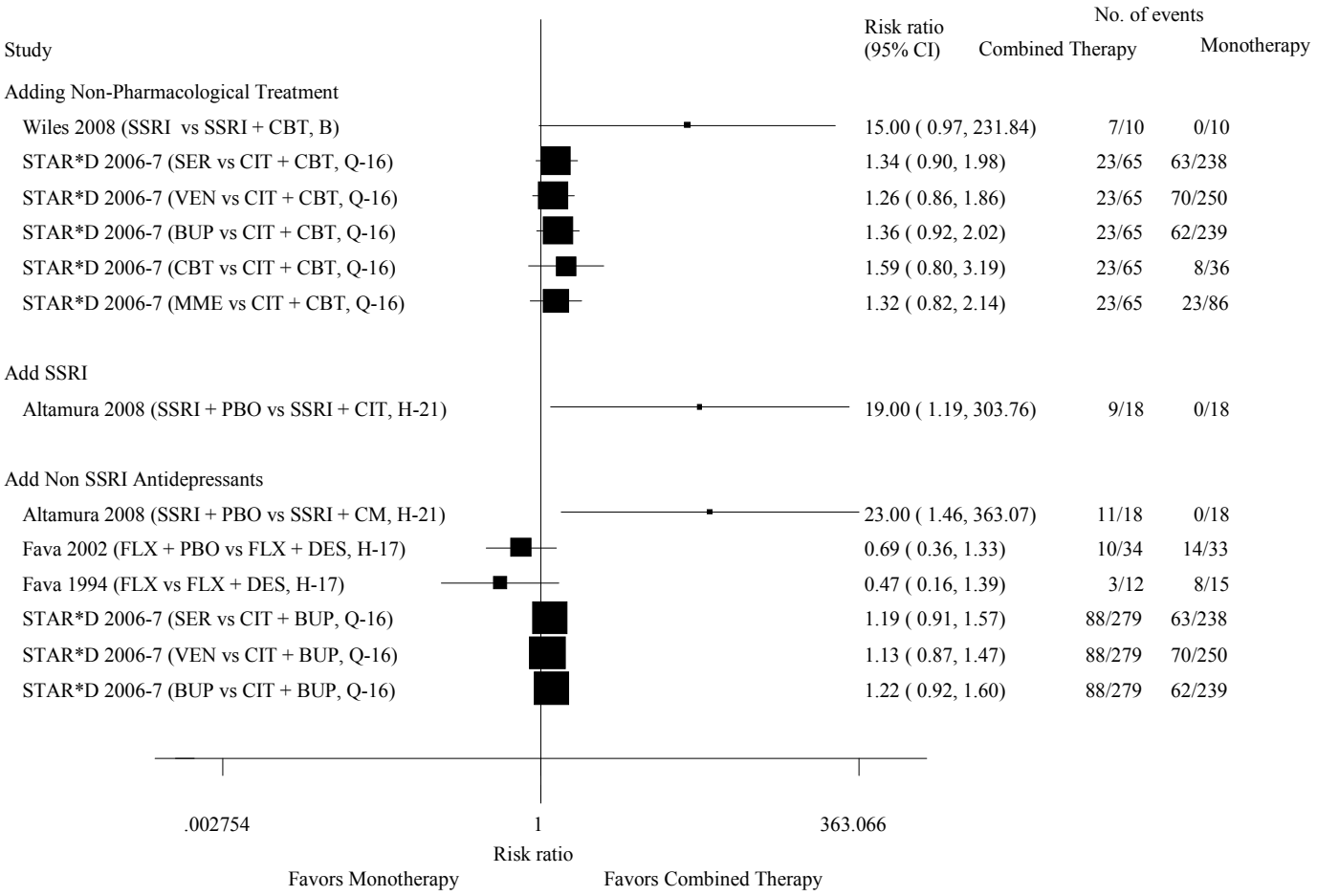


Figure 3.7.Forest plot of monotherapy versus combined therapy for augmenting agents for the outcome of remission

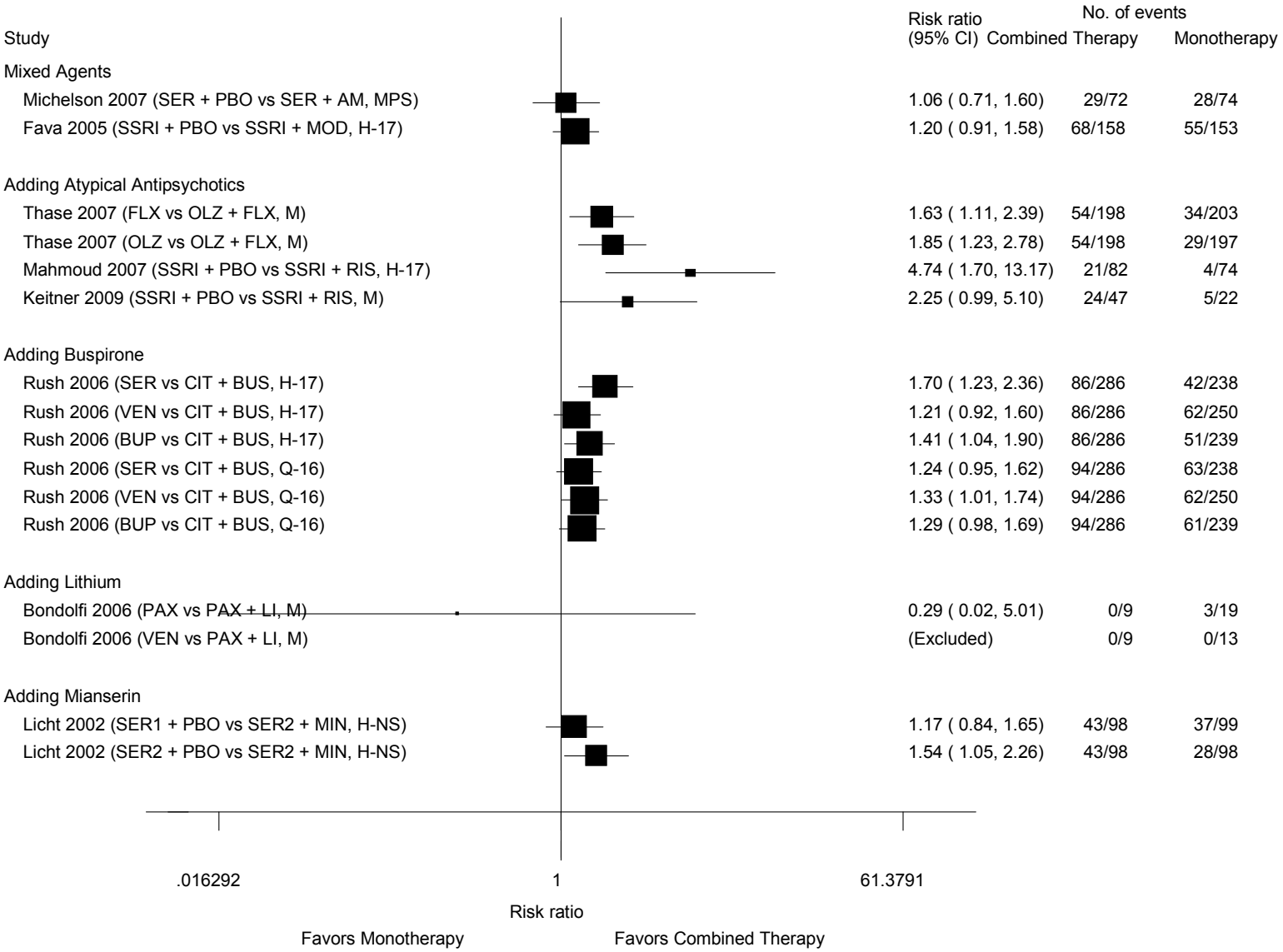
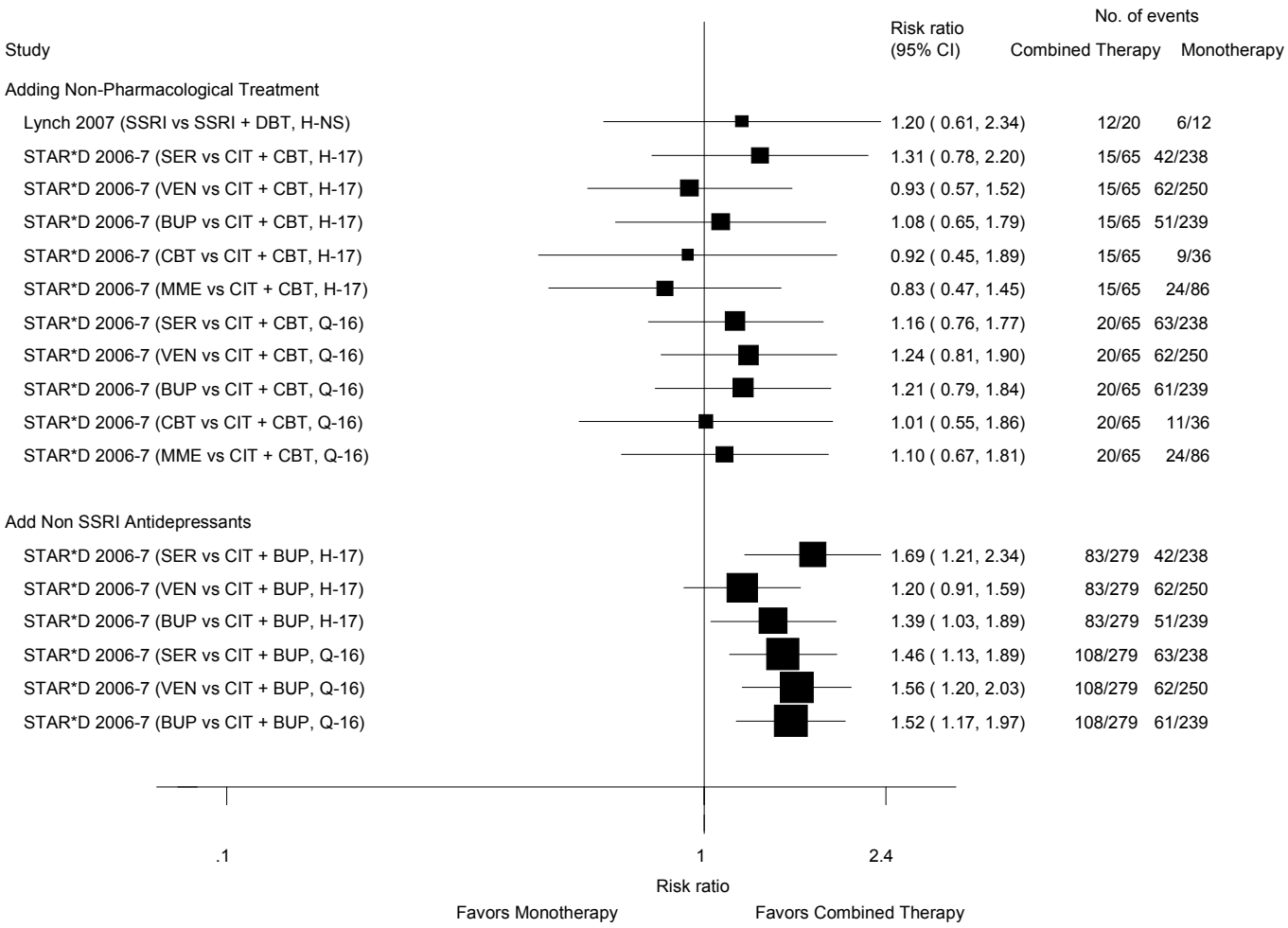


Figure 3.8 Forest plots for monotherapy versus combined therapies for all non-augmenting interventions for the outcome of remission



Combined Therapy versus Combined therapy interventions

There were six studies[113,116,124,135,136,141] for which there were treatment arms that compared combination therapies to each other. All studies were RCT with the exception of one study which did not randomize subjects and the STAR*D study.[142] The STAR*D cohort [113,116] for level 2 subjects, evaluated three combined therapy interventions and only these arms (citalopram plus CBT with two combined drug therapy interventions) are compared in this section. Two studies study[136,142] compared different doses of the same combination drug therapies.

In total there were 832 participants in the treatment arms evaluating combined interventions and the sample sizes varied from 11[141] to 650 participants.[113,116] The sample sizes per treatment arm varied from five subjects[142] to 286 subjects.[116] One study[49,113] exceeded a total sample size of 101 and two studies[135,142] had less than 30 subjects.

Overview of Study PICOT Characteristics

Population. The proportion of women in the sample varied from 47 percent[124,136] between 50 and 62 percent[116,135,142] and greater than 70 percent.[131,132] Racial composition was not reported in four studies;[124,131,132,135,142] two studies reporting ethnicity, had approximately 78 percent[113,116] and over 90 percent[136] of the participants of the white race. Mean age of study subjects varied from 40 to 44 years in four studies,[113,116,124,135,136] and age ranges from 37 to 59 years,[142] and 51 to 58 years[131,132] in the remaining studies.

Inadequate response. Table 11 shows the manner in which failure to an SSRI had been established. Three studies[131,132,135,142]determined failure retrospectively and study subjects were currently on same SSRI prior to the switch to the new intervention. In the three studies that determined inadequate response prospectively, fluoxetine,[143] citalopram,[113,116] and sertraline,[136] were the SSRI for which failure was established. No study evaluated subjects specifically for prospective failed response to escitalopram, paroxetine or fluvoxatine alone.

Table 11. Method of establishing failure to SSRI in studies comparing combined therapies to combination therapies

	COMBINED THERAPIES			
Determining inadequate response	**Add augmentor**	**Add other SSRI**	**Add non-SSRI AD**	**Add non-pharm**
Prospective				
Citalopram	Trivedi[116]*		Trivedi[116]*	Thase[113]*
Fluoxetine	Fava[124]		Fava[124]	
Sertraline	Dunner[136]			
Any SSRI				
Retrospective				
Currently on SSRI or other antidepressant	Fava[135] Dinan[142]	Altamura[131,132]	Altamura[131] Fava[135]	

Mental Health History. Table 12 shows that five studies used the HAM-D 17 or 21 item instruments to evaluate baseline severity; one study did not report baseline scores.[142] It is notable that several studies[49,113,116] included patients of mild to moderate severity based on the HAM-D criteria, while others included patients with marked depression. The number of previous depressive episodes were reported as a median of seven to eight (range 12 to 15) in the STAR*D cohort[113,116] and not reported in five studies.[124,131,132,135,136,141]

Table 12. Distribution of baseline scores for primary outcomes as a proxy for severity of MDD

	Baseline Scores				
Disease Specific Scale	10-14	15- 19	20 - 25	26 to 30	>31
MADRS					
BDI					
HAM-D not specified					
HAM-D-31					
HAM-D-24					
HAM-D-21			Altamura[131] Altamura[132]		
HAM-D-17		Trivedi[116] Rush[49] Thase[113]	Fava[135] Fava[124] Dunner[136]		
QUIDS-SR16	Trivedi[116] Rush[49] Thase[113]				
Other					

Note that one study[142] did not provide baseline scores and some studies provided scores for more than one instrument. No study in this grouping reported baseline use of complementary and alternative medicines (CAM) at baseline or endpoint.

Intervention. All but one study[142] employed an RCT design and the STAR*D is considered a CCT. The STAR*D cohort [104,105,112-114,116] for level 2 subjects, evaluated three combined therapy interventions and only these arms are compared in this section. Two studies[142] compared two doses of the same combination therapy. Table 13 shows the duration of the study intervention; two studies evaluated combined therapy for approximately one week;[131,132,142] the remaining studies varied treatment length from 4 to 12 weeks.

Table 13 details the length of the run-in and treatment phases for all studies

Length of treatment	2/3 weeks	4/5 weeks	6 weeks	8 weeks	>8 weeks
Prospective failure run-in phase			Dunner[136]	Fava[124]	Trivedi[116] Rush[49] Thase[113]
Prospective failure treatment phase		Fava[124]	Dunner[136]		Trivedi[116] Rush[49] Thase[113]
Retrospective failure studies	Dinan[142]# Altamura[131] Altamura[132]##	Fava[135]			

# Indicates treatment was for one week.
Indicates treatment was for 5 days.

Table 14 shows the types of combination therapies evaluated in these six therapies. Two studies included an arm evaluating the non-SSRI desipramine[124,135] and one each evaluating clomipramine[131,132] and buproprion.[116] The augmenting agents used in these studies included buspirone, lithium, and ziprasidone. Two studies[136,142] compared different doses of the same combination studies involving sertraline with either lithium and ziprasidone. The doses for both lithium (400-800mg) and ziprasidone (60mg and 80mg) are in the low to moderate range. It is unlikely that lithium at 400mg/d would result in therapeutic blood levels, but low doses of lithium have been commonly employed in augmentation trials. The STAR*D cohort compared two drug combination therapies with citalopram or CBT.

Table 14. Combined therapy versus combined therapy studies showing the comparison and treatment interventions grouped by type of intervention

Study	Combined therapy 1	Combined therapy 2
Altamura[131] 2008 Altamura[132] 2008	SSRI + Citalopram (intravenous)	SSRI + Clomipramine (intravenous)
Rush[49] 2006 Trivedi[116] 2006 Thase[113] 2007	Citalopram + Buproprion	Citalopram + Buspirone
Add Augmenting Agent		
Dinan[142] 1993	Sertraline + Lithium 400mg	Sertraline + Lithium 800mg
Dunner[136] 2007	Sertraline + Ziprasidone 60mg/d	Sertraline + Ziprasidone 80mg/d
Fava[124] 2002	Fluoxetine + Desipramine	Fluoxetine + Lithium
Fava[135] 1994	Fluoxetine + Desipramine	Fluoxetine + Lithium
Adding Non-pharmacological Treatment		
Thase[113] 2007 Rush[49] 2006 Trivedi[116] 2006	Citalopram + Buspirone Citalopram + Buproprion	Citalopram + CBT

*indicates that comparison arm is not the SSRI prior to the switch.

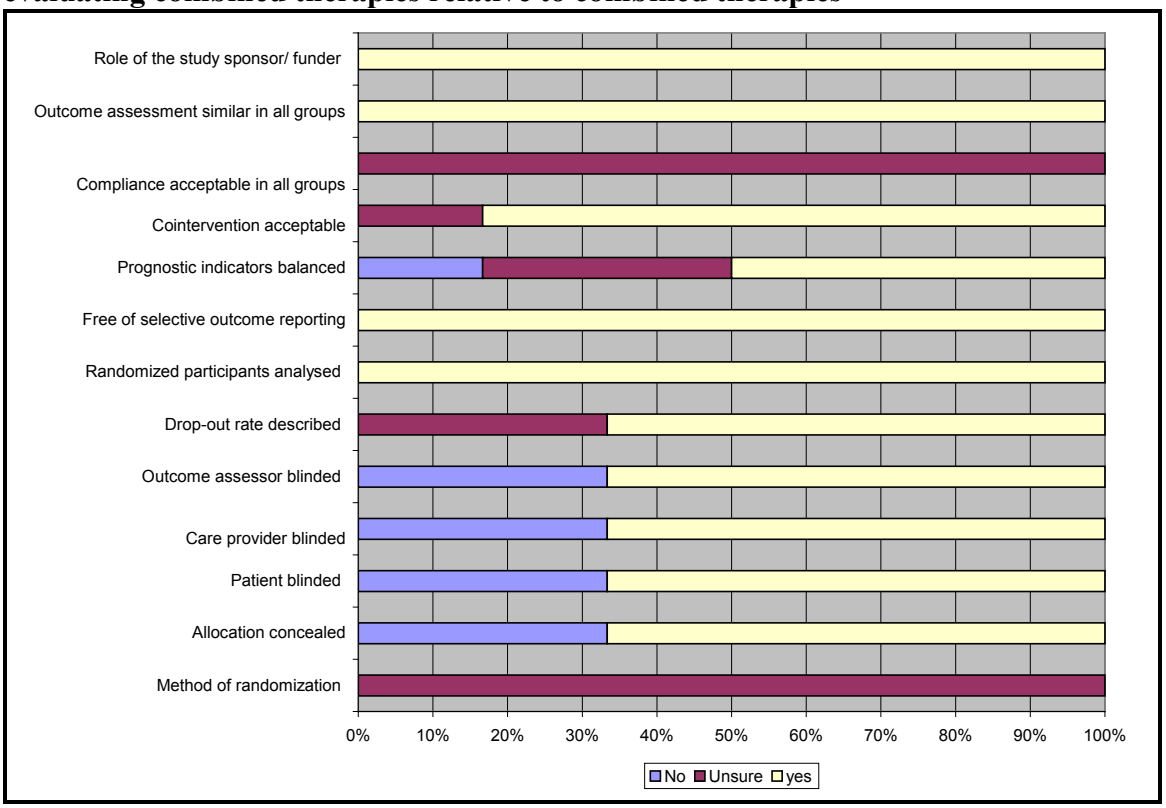
Outcome. A single study[113,116] specified that remission was the primary outcome. All other studies indicated that the change or endpoint score was the primary outcome.

Setting. The six studies were conducted in the UK[142] Italy,[131,132] Canada,[136]and the United States (4 studies).[104,105,112-114,116,124,135,136] All studies included subjects in outpatient psychiatric or outpatient primary care.[124]

Risk of Bias

Figure 3.9 shows that studies evaluating combined therapies were at high risk of bias for randomization, reporting compliance, and balancing prognostic indicators. The role of the funder was clarified in all studies and funding for the studies came from non-industry sources in three studies,[116,124,135] industry in one study,[136] and two did not reported the source.[131,132,142] Overall these studies would be categorized as having moderate level of risk of bias. None of the studies employed a washout phase or monitored compliance of subjects.

Figure 3.9. Percent of studies achieving risk of bias using the McHarm criteria for studies evaluating combined therapies relative to combined therapies



Efficacy of Combined Therapy vs. Combined Therapy

Tables 15 & 16 report the rates of response and remission for comparison of combined treatments. Figure 3.10 illustrates that when the combination of citalopram plus buspirone was compared against the combination of citalopram and CBT, there was a non-significant pattern favoring the combination of medications in the STAR*D trial. There appeared to be no advantage of one combination therapy relative to any other.

Table 15. Summary of reported rates of response and remission for studies comparing monotherapy treatment to combined therapy

Study	Duration (weeks)	Rating Scale	N*	Comparison and Dose (mg/d)	Response N (%)	P value	Remission N (%)	P value
ADD SSRI								
Altamura[131,132] 2008	5 day	HAMD-21	18	SSRI + PBO (saline)	0 (0)	<0.0001		
			18	SSRI + CIT 10mg in 250ml of saline	9 (50)			
ADD NON-SSRI Antidepressants								
Altamura[131,132] 2008	5 day	HAMD-21	18	SSRI + PBO (saline)	0 (0)	<0.0001		
			18	SSRI + CM 25mg in 250ml of saline (intravenous)	11 (61.1)			
Fava[124] 2002	12	HAMD-17	33	FLX 40-60mg/d + PBO	14 (42.4)	0.2		
			34	FLX 20mg/d, DES 25-50mg/d	10 (29.4)			
Fava[135] 1994	4	HAMD-17	15	FLX 40-60mg/d	8 (53)	0.24		
			12	FLX 20mg + DES 25-50mg/d	3 (25)			
Rush[49] 2006	12	HAMD-17 QUID-SR-16*	238	SER 50-200mg/d	63 (26.7)*	NR	42 (17.6) 63 (6.4)*	
			250	VEN 37.5-375mg/d	70 (28.2)*		62 (24.8) 62 (24.8)*	
			239	BUP 150-400mg/d	62 (26.1)*		51 (21.3) 61 (25.5)*	
Trivedi[116] 2006	12	HAMD-17 QUID-SR-16*	279	CIT + BUP, 200-400mg/d	88 (31.8)*		83 (29.7) 108 (38.7)*	0.93 0.16*
Preskorn[126] 2008	6	HAMD-NS	15	PAX 40mg + PBO	3 (20)	<0.10		
				PAX 40mg + CP-101,606 infusion/duration to 1.5 hours and the dose to 0.5mg/kg per hour	12 (80)			
George[127] 2008	8	HAMD-17	10	SSRI + PBO	1 (10)	0.15		
			11	SSRI + ME, 5mg PO/d	5 (45.4)			
AUGMENTING AGENTS								
Michelson[119] 2007	8	MPS	74	SER 100mg/d + PBO			28 (37.8)	0.865
			72	SER 100mg/d + AM 40mg/d			29 (40.3)	
Shapira[128] 2006	4	HAMD-21	9	SSRI + PBO	7 (9)	0.02		
			11	SSRI + PI	2 (11)			
Seidman[129] 2005	6	HAMD-24	13	SSRI + PBO volume matched, (injection)	3 (23.1)	0.226		
			13	SSRI + TE 200-600mg/d	7 (53.8)			

Table 15. Summary of reported rates of response and remission for studies comparing monotherapy treatment to combined therapy

Study	Duration (weeks)	Rating Scale	N*	Comparison and Dose (mg/d)	Response N (%)	P value	Remission N (%)	P value
Berman[77] 2007	8	MADRS	176	New SSRI + PBO	NR for SSRI subgrup		NR for SSRI subgrup	
			182	New SSRI + ARI, 5-15	NR for SSRI subgroup		NR for SSRI subgroup	
Marcus[78] 2008	8	MADRS	191	SSRI + PBO	NR for SSRI subgrup		NR for SSRI subgrup	
			190	SSRI + ARI 5-20mg/d (5-11mg ?)	NR for SSRI subgroup		NR for SSRI subgroup	
Fava[123] 2005	8	HAMD- 17 CGI-1*	153	SSRI + PLB	48 (32)*	>0.09	55 (36)	0.2
			158	SSRI + MOD 100-200	64 (41)*		68 (44)	
Nemets[134] 1999	4	HAM-D 24	18	SSRI original dose + PBO	NR			
			18	IN 12gm/d, SSRI original dose;	NR			
Dunner[136] 2007	8	MADRS	20	SER 100-200mg/d	4 (19)	NS		
			21	SER 100-200 mg/d + ZI 40-80mg/d	6 (32)			
			19	SER 100-200 mg/d + ZI 80-160mg/d	2 (10)			
Adding Atypical Antipsychotics								
Thase[108] 2007	8	MADRS	203	FLX 50mg/d			34 (16.7)	
			197	OLZ 6-18mg/d			29 (14.7)	0.012
			198	OLAN 6-18mg/d + FLX 50mg/d			24 (27.3)	0.003
Shelton[109] 2001	8	MADRS	8	FLX 20-60mg/d	1 (10)	0.006 vs. com		
			8	OLZ 5-20mg/d	0 (0)	0.003 vs. com		
			10	OLZ 5-20mg/d, FLX 20-60mg/d	6 (60)	0.007		
Mahmoud[76] 2007	6	HAMD-17	74	SSRI + PBO	18 (24.3)	NR	4 (5)	NR
			82	SSRI + RIS 0.25-1mg/d	37 (45.70		21 (25)	
Keitner[75] 2009	4	MADRS	22	SSRI dose maintained + PBO			5 (22.7)	0.011
			47	RIS 0.5-3mg/d + antidepressant dose maintained			24 (51)	
Adding BUS								
Rush[49] 2006	12	HAMD-17 QUID-SR-16*	238	SER 50-200mg/d	63 (26.7)*	NR	42 (17.6) 63 (6.4)*	
			250	VEN 37.5-375mg/d	70 (28.2)*		62 (24.8) 62 (24.8)*	

Table 15. Summary of reported rates of response and remission for studies comparing monotherapy treatment to combined therapy

Study	Duration (weeks)	Rating Scale	N*	Comparison and Dose (mg/d)	Response N (%)	P value	Remission N (%)	P value
			239	BUP 150-400mg/d	62 (26.1)*		51 (21.3) 61 (25.5)*	
Trivedi[116] 2006	12	HAMD-17 QUID-SR-16*	286	CIT + BUS, 200-400mg/d	77 (27)*		86 (30.1) 94 (32.9)*	0.93 0.16*
Appelberg[125] 2001	6	MADRS	51	CIT 40mg/d/FLX 35.4mg/d + PBO	16 (31)	0.034		
			51	CIT 40mg/d/FLX 35.4mg/d + BUSP 35-47mg/d	17 (33)			
Landén[120] 1998	4	CGI-S, CGI-I	60	CIT 46.1mg/d or PARO 39.8mg/d + PBO	28 (46.7)	NS		
			57	CIT 46.1mg/d or PARO 39.8mg/d + BUSP 49mg/d	29 (50.9)			
Adding Li								
Fava[135] 1994	4	HAMD-17	15	FLX 40-60mg/d	8 (53)	0.24		
			14	FLX 20mg/d + LI 300-600mg/d	4 (29)			
Fava[124] 2002	12	HAMD-17	33	FLX 40-60mg/d + PBO	14 (42.4)	0.2		
			34	FLX 20mg/d, LI 300-600mg/d	8 (23.5)			
Baumann[130] 1996	4	HAMD-21	14	CIT 40-60mg/d	2 (14)	0.05		
			10	CIT 40-60mg/d, LI 800mg/d;	6 (60)			
Bondolfi[91] 2006	4	MADRS	19	PARO 40mg/d	2 (10.5)	NR	3 (15.7)	NR
			9	VEN 150mg/d	0 (0)		0 (0)	
			13	PARO 30mg/d + LI	1 (7.8)		0 (0)	
Adding PIDOLOL								
Perry[133] 2004	6	HAMD	17	SSRI + PBO FLX 20-60mg, PARO20-40mg, SER 50mg	6 (35.2)	1		
			21	SSRI + PI Total = only SSRI doses given, PI dose not reported; Group 1 = FLX 20-60mg, PARO 20mg, SER 150-200mg	5 (23.8)			
Sokolski[118] 2004	4	HAMD	5	PARO 40mg/d + PBO	NR	0.001		
			4	PARO 40mg/d + PI 7.5mg/d	3 (75)			
Adding MIN								
Licht[110] 2002	6	HAMD-NS	99	SER 100mg/d + PBO	69 (70)	0.64	37 (38)	0.38
			98	SER 200mg/d + PBO	54 (64)	0.03	28 (29)	0.19

Table 15. Summary of reported rates of response and remission for studies comparing monotherapy treatment to combined therapy

Study	Duration (weeks)	Rating Scale	N*	Comparison and Dose (mg/d)	Response N (%)	P value	Remission N (%)	P value
			98	SER 100mg/d + MIN	66 (67)		43 (44)	
Ferreri[111] 2001	6	HAMD 17	38	FLX 20mg/d	14 (37)	0.1	14 (36)	0.06
			34	MIN 60mg/d	16 (48.5)		6 (18)	
			32	FLX 20mg/d + MIN 60-60mg/d	20 (62.5)		14 (44)	
Adding Non-Pharmacological								
Carta[81] 2008	32	WHOQOL-Bref Pych;	10	SSRI	NR		NR	
			20	SSRI + Exercise	NR		NR	
Lynch[117] 2007	54	HAMD-NS	12	SSRI			6 (50)	NR
			20	SSRI + Dialectical Behavioural Therapy			12 (60)	
Wiles[74] 2008	16	BDI	9	SSRI	0			
			14	SSRI + CBT	8 (56)			
Rush[49] 2006	12	HAMD-17 QUID-SR-16*	238	SER 50-200mg/d	63 (26.7)*	NR	42 (17.6) 63 (6.4)*	
			250	VEN 37.5-375mg/d	70 (28.2)*		62 (24.8) 62 (24.8)*	
			239	BUP 150-400mg/d	62 (26.1)*		51 (21.3) 61 (25.5)*	
			36	Medications Monotherapy	8 (22.2)		9 (25) 11 (30.5)	
			86	CBT	23 (26.7)		24 (27.9) 24 (27.9)	
			65	CIT + CBT	23 (35.4)		15 (23) 20 (20.7)	
			250	VEN 37.5-375mg/d	70 (28.2)*		62 (24.8) 62 (24.8)*	
			239	BUP 150-400mg/d	62 (26.1)*		51 (21.3) 61 (25.5)*	
			36	Medications Monotherapy	8 (22.2)		9 (25) 11 (30.5)	
			86	CBT	23 (26.7)		24 (27.9) 24 (27.9)	
			65	CIT + CBT	23 (35.4)		15 (23) 20 (20.7)	

ABBR: BUP = , CIT=, CM=, DES=, FLX=, HAMD=Hamilton Depression Scale, ME=, N=sample size, p = Probability, PAX=, PBO=, PI=, SSRI=, QUID-SR=, TE=, VEN=

Table 16. Summary of reported rates of response and remission for studies evaluating combined therapy to other combined therapy treatments

Study	Duration (weeks)	Rating Scale	N*	Comparison and Dose (mg/d)	Response N (%)	P value	Remission N (%)	P value
ADD Non-SSRI								
Altamura[131,132] 2008	5 days	HAMD-21	18	SSRI+CIT 10mg in 250ml of saline	9 (50)			
			18	SSRI+ CM 25mg in 250ml of saline (intravenous)	11 (61.1)			
Trivedi[116] 2006	12	HAMD-17 QUID-SR-16*	286	CIT +BUS 15-60mg/d	77*		86 (30.1) 94 (32.9)*	0.93 0.16*
			279	CIT +BUP, 200-400mg/d	62 (22.2)*		83 (29.7) 108 (38.7)*	
AUGMENTING AGENTS								
Dinan[142] 1993	1	HAMD-NS	6	SER 100-200mg/d + LI 400mg/d	4			
			5	SER 100-200mg/d + LI 800mg/d	3			
Dunner[136] 2007	8	MADRS	21	SER 100-200 mg/d + ZI 40-80mg/d	6 (32)			
			19	SER 100-200 mg/d + ZI 80-160mg/d	2 (10)			
Fava[124] 2002	12	HAMD-17	34	FLX 40-60mg/d + placebo DES	10 (29.4)			
			34	FLX 20mg/d, LI 300-600mg/d	8 (23.5)			
Fava[135] 1994	4	HAMD-17	12	FLX 20mg + DES 25-50mg/d	3 (25)			
			14	FLX 20mg/d + LI 300-600mg/d	4 (29)			
Adding Non-Pharmacological								
Trivedi[116] 2006 Thase[113] 2007	12	HAMD-17 QUID-SR-16*	286	CIT + BUS 15-60mg/d	77 (27)*		83 (29.7) 94 (32.8)*	0.93
			239	CIT + BUP, 200-400mg/d	62 (26.1)*		94 (39.3)	
Thase[113] 2007	12	HAMD-17 QUID-SR-16*	117	Medications Combined	33 (28.2)*		39 (33.3) 39 (33.3)*	
			65	CIT + CBT	23 (35.4)*		15 (23.1 20 (30.8)*	

Figure 3.10. Forest plots of the combined versus combined therapies for the outcome of response

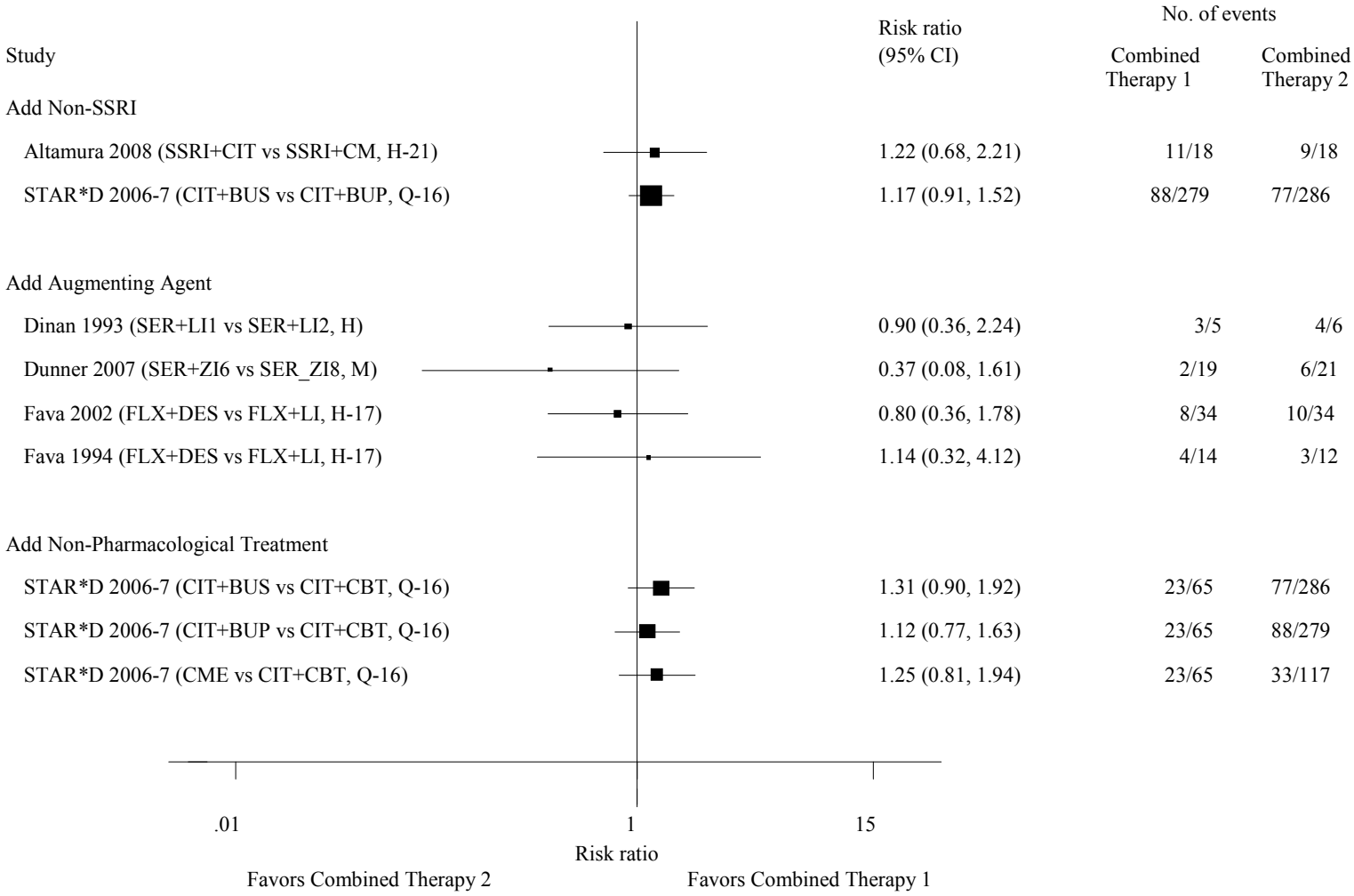
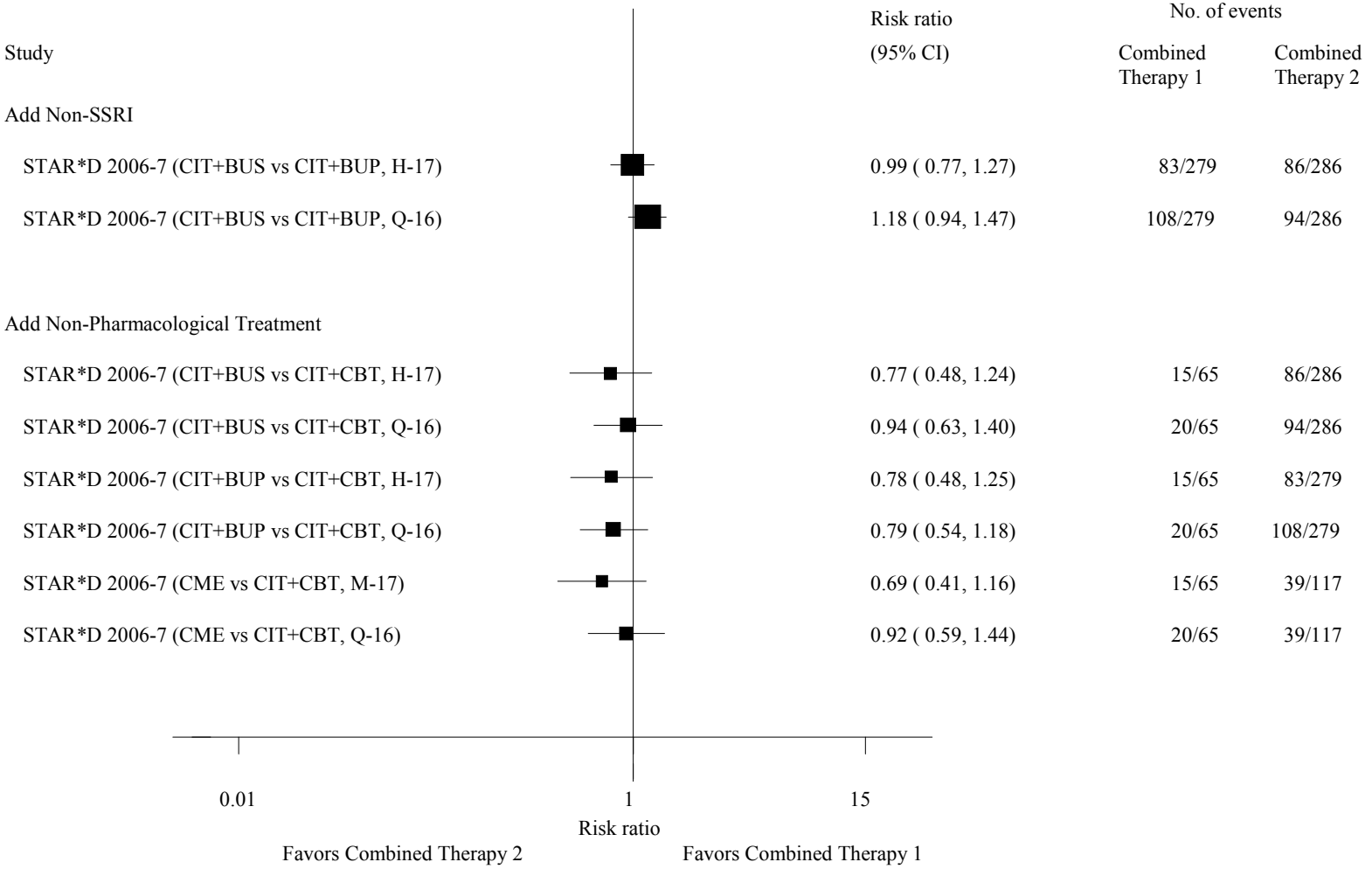


Figure 3.11. Forest plot of combined versus combined therapies for the outcome of remission



Interventions in Patients with Subsyndromal Depression or Dysthymia

Overview of Study PICOT Characteristics: Subsyndromal Depression

Population. A single study[100] evaluated patients described as those "with residual symptoms of a depressive disorder" and characterized by a score greater than 7 but less than 10 on the HAM-D-21 items. These subjects were classified as having subsyndromal depression following an acute episode. Seventy percent of the subjects were women and ethnicity was not reported. Mean age was 39 years.

Inadequate response.: There were no specific criteria reported for previous failure to paroxetine other than having residual symptoms and having been treated for 42 to 300 days.

Mental Health History. Failure to paroxetine was determined prospectively over a 4 week period to paroxetine. The subjects' failures to the current treatments were retrospective but the manner of determining this was not reported. Similarly, the history of any previous inadequate responses to treatment or length of the current episode was not reported

Intervention and Comparators. In this study subjects who had residual symptoms while on paroxetine were randomized to a continuation of paroxetine (20-40mg/d) or switched to mirtazapine (15 to 30mg/ day) for an average of 36 days.

Outcomes. The primary outcomes in this study[100]were rated on the HAM-D-21. Changes in metabolic rate values and changes in the Arizona Sexual Experience Scale (ASEX) score were also reported as not different between groups.

Setting. This study was conducted in the Czech Republic and the setting from which patients were recruited was not reported.

Risk of Bias: Subsyndromal Depression

In this study, the type of randomization process and the degree of compliance was not clearly reported; all other categories were acceptable.

Efficacy of treatment: Subsyndromal Depression

The findings of this study do not report differences between groups; rather it reported that 70 percent of subjects had a positive effect on residual symptoms but no mean change scores were given. Differences between groups were shown on the scale for sexual functioning in favor of mirtazapine ($p = 0.004$) at four weeks.

Overview of Study PICOT Characteristics: Dysthymia

Population. One study[101] evaluated subjects with dysthymia as diagnosed by the DSM-IV structured clinical interview and with a score of 12 or more on the HAM-D-21 scale. Subjects with MDD or other types of depression (partial remission from depression) were excluded. Sixty eight percent of the sample were women and the mean age was 42 years. Ethnicity was not reported.

Inadequate response. Subjects were not excluded because of failures other than the current response to paroxetine. The number of previous episodes of failure to treatment were not reported but the mean duration of the depression was approximately 12 years with an onset approximately at 29 years of age.

Mental Health History. The subjects' failure to the current treatment was retrospective but the manner of determining this was not reported. Similarly, the history of any previous inadequate responses to treatment or length of the current episode was not reported.

Intervention and Comparators
Subjects were randomized to either paroxetine (40mg/d) or paroxetine (20mg/d) and amisulpride (50mg/d).

Outcomes
The primary outcome for the study was response (defined as 50 percent change from baseline) for the HAM-D (type not specified) and a score of one or two on the CGI-2. Remission was secondary outcome and was not explicitly defined but assumed to be defined as a score on the HAM-D.

Setting
The study was conducted in Italy and subjects were recruited from outpatient settings.

Risk of Bias Dysthymia

This paper was at low risk of bias and there was only uncertainty around the role of the study sponsor.

Efficacy of treatment in Dysthymia

Fifty four percent of subjects on paroxetine alone and 56 percent in the combined group achieved response (50 percent change) on the HAM-D-NS. Remission was defined as a score of seven or less and those achieving remission were 32 percent for paroxetine alone and 44 percent for the combination treatment group. Neither response nor remission was shown to be statistically different between the treatment groups.

Summary of Findings for Adults and Strength of Evidence

We applied the criteria for grading the strength of evidence for the studies for MDD only. Single trials for patients with dysthymia and subsyndromal depression were not considered in this grading.

Monotherapy vs. Monotherapy and Combined Therapy vs. Combined Therapy in MDD

The grading of the strength of evidence (SOE) for adults with MDD who have failed to respond to an SSRI is detailed in Appendix C Table 1. With respect to monotherapy compared to monotherapy interventions, we grouped all treatment approaches together given the small number of studies and the varying drugs and CBT. Overall, there were study limitations related to randomization, and the indirectness of the populations and interventions that contributed to the rating of a low SOE. The confidence intervals were generally small and the effect sizes of similar magnitude (all non significant); all comparisons indicated that there was no advantage of any one monotherapy over another. This was also the case for the STAR*D monotherapy comparisons despite the relatively large sample sizes per group relative to other studies. A similar rating for the SOE was given to the studies that compared combined therapies relative to other combined therapies; the STAR*D study was the single study in this group reporting the outcome of remission. These studies were consistent in that the relative risks were generally of the same magnitude and the effect sizes showed that no one combined therapy was better than any other comparison; given the study limitations, we rated these as low quality SOE.

Monotherapy versus Combined Therapies in MDD

The SOE ratings for the studies comparing monotherapies relative to combined therapies is detailed in Appendix C Table 2. The greatest number of studies compared monotherapy relative to adding augmenting agents. We considered these augmenting studies both as a group and as subgroups related to the number of studies evaluating specific agents; there were four subgroups we considered with respect to specific classes of agents and these included, atypical antipyschotics, buspirone, lithium, mianserin and then all other agents were categorized as a single group for SOE rating. When we considered all studies with augmenting agents (N = 10) as a single group, we rated the studies evaluating monotherapies relative to adding augmenting agents as insufficient SOE; the large number of treatment agents, differing treatment intervals and population characteristics, and the wide range of sample sizes contributed to this grading of insufficient.

When considering atypical antipsychotics alone, a SOE rating of moderate for the outcome of remission was given; there was a consistent effect favoring combined treatment and these studies had relative larger sample sizes. We note that two studies[75,76] showed large confidence intervals and this may be related to the "subgroup" data specific to failed SSRI group that we requested from the study authors. The original study data included larger sample sizes as subjects with failed response to non-SSRI drugs were included. A SOE grade of low was given for the outcome of response as one study with a very small sample size had large variability.

We separated the buspirone studies into those that switched to different antidepressant monotherapy versus those that added buspirone to the current SSRI. The SOE was graded as

insufficient for the later category (adding) as the studies were deemed to have a greater number of study limitations relative to the STAR*D trial that evaluated switching to new monotherapies. The STAR*D trial again showed no advantage to the combined buspirone combined therapy relative to the different monotherapies.

The remaining groupings for augmenting agents for lithium, mianserin, and "other agents" were all graded as insufficient SOE due to the small sample sizes, and significant study limitations. It is difficult to determine any level of confidence in the effects of these agents despite the fact that none were shown to have any advantage over the comparator monotherapy.

We grouped all studies that maintained the current SSRI and then compared this treatment arm to one where a different SSRI, non-SSRI, or nonpharmacological treatment was added. This group of studies was rated as low for the outcome of response because of the differing agents and small sample sizes accounted for this rating; for the outcome of remission, a grading of insufficient was given, as the study limitations were significant. There were two studies that compared switching from the current SSRI to a new monotherapy treatment and then comparing this with the new agent combined with any other drug. The studies evaluating switching to a new antidepressant and then adding ariprazole would have been included in this group, had we been able to acquire the rates of response and remission for the SSRI failed group. For the two studies that did provide these outcomes, one study[91] had wide confidence intervals and effect size because of the small sample size; the other study was the STAR*D cohort and had multiple treatment arms and comparisons. The evidence is graded as low and the findings suggest no relative advantage to switching to a new drug or CBT relative to a group that combined buspirone or buproprion.

Adolescents

Overview of Study PICOT Characteristics in studies with Adolescents

There were two studies evaluating adolescents who had not responded to previous SSRI treatment but only one[17] could be extracted. A second trial[97-99] did have "Phase II" subjects (those who had had inadequate response) and two of the three study arms were eligible for this review (medication or CBT). Proportions of subjects who reached this stage were reported and contact with the authors confirmed that data is not currently available for Phase II subjects.

Population. In the Treatment for Resistant Depression in Adolescents (TORDIA) study,[17,95,96] the majority of the sample (68 to 72 percent) were female adolescents from age 12 to 18; the average age was 16 years (SD 1.6). The sample was predominately (> 80 percent) white.

Inadequate response. Failure to an adequate dose of an SSRI was established retrospectively in this study but subjects were currently on fluoxetine. Subjects who had previously failed two or more adequate trials of an SSRI, had a history of non-response to venlaxafine, or non response to CBT (greater than or equal to 7 sessions) were excluded from the study. Potential participants who were currently receiving CBT or were on other medications with psychoactive properties were also excluded.

Mental Health History. For approximately 74 percent of the sample, this was their first episode of MDD. The mean duration of the current episode varied from 21 to 24 months. Approximately one quarter of the sample had a history of suicide attempts (varying from 21 to 27 percent). The level of co-morbidity was significant in this group and approximated 36 percent for anxiety disorder, 21 to 24 percent post traumatic stress disorder, 14 to 18 percent for attention deficit hyperactivity disorder, and 27 to 32 percent with dysthymia. However, there were no differences in rates of co-morbidity between the four treatment groups. Baseline severity scores varied from 19 to 22 on the Beck Depression Inventory (BDI) and 58 to 60 on the Children's Depression Rating Scale- Revised (CDRS-R).

Intervention and Comparators. Study subjects were randomized to four treatment arms that included venlafaxine alone (up to 150mg/d), venlafaxine combined with CBT, citalopram, fluoxetine, or paroxetine (up to 40mg/d for all SSRI) alone or with CBT. CBT consisted of up to 12 (60-90 minute) sessions and one quarter to one half consisted of sessions with the family. The reported mean number of sessions was 8.3 across treatment groups. Subjects were tapered off the initial SSRI. All participants received family psychoeducation which consisted of providing information about depression, adverse effects ad coping with mood disorders. The treatment interval was 12 weeks.

Outcome. This study had two primary outcomes based on "adequate clinical response" defined as a score of two or less on the Clinical Global Impression Improvement subscale and a 50 percent improvement on the CDRS-R reported both as a percent achieved and as trajectory over time.

Setting. This study was conducted in the U.S. and adolescents were recruited from clinical sources (80 percent) and advertisements (20 percent).

Risk of Bias in Studies with Adolescents

There was some uncertainty with regards to the method of allocation concealment, blinding of the outcome assessor, but this trial maintained low risk of bias in all other aspects of the TORDIA trial. A washout period for subjects on SSRI other than fluoxetine was undertaken for 2 weeks prior to switching to the new intervention. The method of assessing compliance with the treatment was not reported, but the proportion of subjects who did not comply was reported. Overall, the TORDIA trial has a low risk of bias.

Treatment fidelity for the CBT was well detailed and approximately 94 percent of reviewed tapes were found to be acceptable by on site supervisors and by an external consultant.

Efficacy of treatment in Adolescents

Treatment results are presented as SSRI (with and without CBT), and venlaxafine (with and without CBT) groups and then the no CBT group (SSRI and venlaxafine combined) or CBT group (SSRI + CBT and venlaxafine + CBT). Study results for response are detailed in Table 17. There were no differences between the medications groups, but there was a statistically

significant difference between the CBT groups in favor of including CBT for all outcomes. The main effect of CBT was consistent even after controlling for a number of baseline severity factors (BDI scores and post-traumatic stress).

Table 17. Results from TORDIA trial for ITT sample

ITT sample	SSRI (N = 168)	Venlaxafine (N = 166)	No CBT (N = 168)	CBT (N = 166)
Response* (%)	79 (47.0)	80 (40.5)	68 (40.5)	91 (54.8)
CGI-I =<2 (%)	86 (51.2)	92 (55.4)	80 (47.6)	98 (59.0)
Change CDRS-R >= 50% (%)	86 (51.8)	86 (51.8)	79 (47.0)	191 (60.8)

* Response defined as "adequate clinical response" on the CGI-I

KQ2. What are the harms of each of the monotherapy or combined therapies among these adults and adolescents? How do the harms compare across different interventions?

KEY MESSAGES
Harms for interventions used in both adults and adolescents with MDD who had failed to respond to SSRI were derived from predominately RCTs that evaluated treatment strategies in this population; no observational studies were eligible. A clear trend for harms was difficult to specify across the differing interventions in adults. Harms were well evaluated in the single study in adolescents.
Reporting and collecting of harms was problematic, particularly for pre-defining harms including serious and severe events and reporting total number of events per group in study with studies with adults. The single study evaluating harms in adolescents provided high quality evidence for harms within this population when receiving pharmacological and psychological treatment.
Severe events and serious events (including suicidality) were inconsistently reported in studies with adult MDD populations.
A limited number of studies undertook statistical evaluation comparing harms between groups.

Harms in Adults with MDD, Dysthymia, and Subsyndromal Depression

From the 37 studies evaluating adults and all but one study included subjects with MDD; two studies evaluated subjects with Subsyndromal Depression[100] and Dysthymia.[101] As noted previously, five studies[89-91,94,102] and seven STAR*D publications[82,83,85-88,92] did not have data that could be extracted. No observational studies with the required patient population and evaluation of harms was eligible for this CER. The summary of harms thus reflects those reported within the eligible studies.

We present the harms evidence for the eligible and extracted studies based on the type of treatment comparisons as follows: 1) monotherapy compared to monotherapy, 2) monotherapy

compared to combined therapy, and 3) combined therapy compared to combined therapy. Some studies evaluated more than two treatment arms, and are included in multiple sections dependent on the drugs used.

Description of studies in reporting Harms in Adults with MDD

Monotherapy versus Monotherapy in Adult MDD

All but one study[91] reported some aspect of safety and tolerability in the six studies having at least one monotherapy treatment arm. None of the studies were specifically designed to compare the effect of harms between different monotherapies.

The method of assessing adverse effects differed greatly among studies with a limited number of studies using standardized methods or scales. Figure 3.12 shows the ratings on the McHarm scale for evaluating risk of bias and reporting within comparative studies. Forty percent of the studies indicated that the harms reported were those that were observed in two or three percent,[115] five percent[110] or 10 percent of subjects;[108] the remaining studies did not specify why the harms reported were included or were unclear (60 percent). None of the studies provided any a priori definitions of the harms, definitions of serious or severe events. Similarly, the mode of how harms were collected or the training of the person collecting them was not specified. Generally, the number of subjects who withdrew were specified per treatment arm; however, the number of specific adverse events per treatment arm were not well specified (42 percent).

Figure 3.12 Ratings of studies evaluating monotherapies using the McHarm criteria for risk of bias and reporting

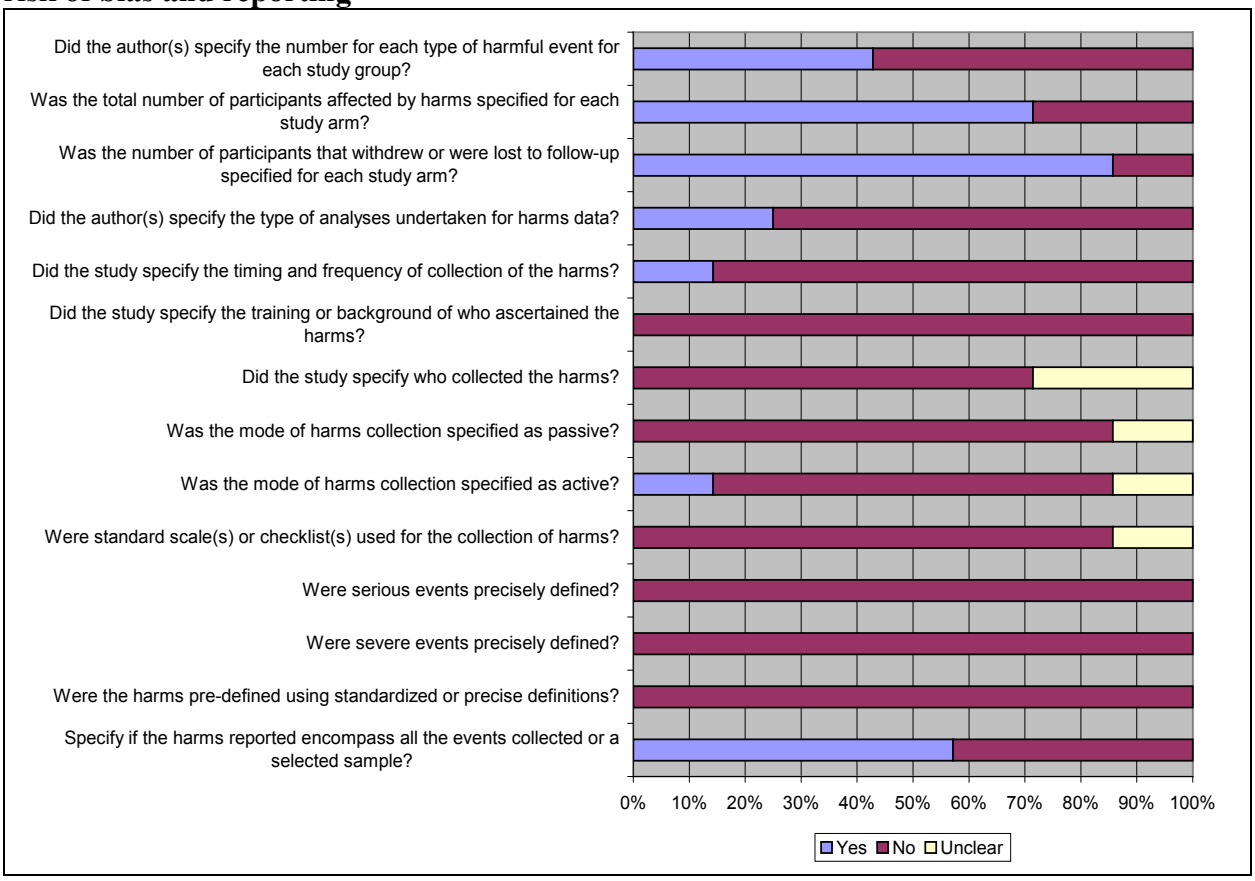


Table 18 shows the rates of reported harms as a function of the treatment arm. We selected seven main categories of harms. The STAR*D cohort reported only the frequency of events as a range from 1 to 100 percent (not specifying the types of events as individual frequencies), and similarly, identified numbers of serious events as having at least one event.[49,113,116] Three studies explicitly identified that no serious events had occurred,[49,113,116] or that suicide events had explicitly not occurred;[49,108,113,116] for the STAR*D trials we are assuming that serious psychiatric events encompasses suicidality. Rates of discontinuation due to adverse events were variable. In studies with open label prospective failure components, the number of patients who had adverse events who did not proceed to the next phase was not consistently reported. In studies with historical failure, the proportion of subjects who had experienced inadequacy due to intolerance because of harms was not detailed.

Two studies reported on both serious and suicide related events.[49,113,115]Other adverse events not reported in Table 18 include dry mouth,[108,110,111,115] dizziness,[111,115] and fatigue;[108,111] increased appetite or weight gain was reported in two studies.[108,109]

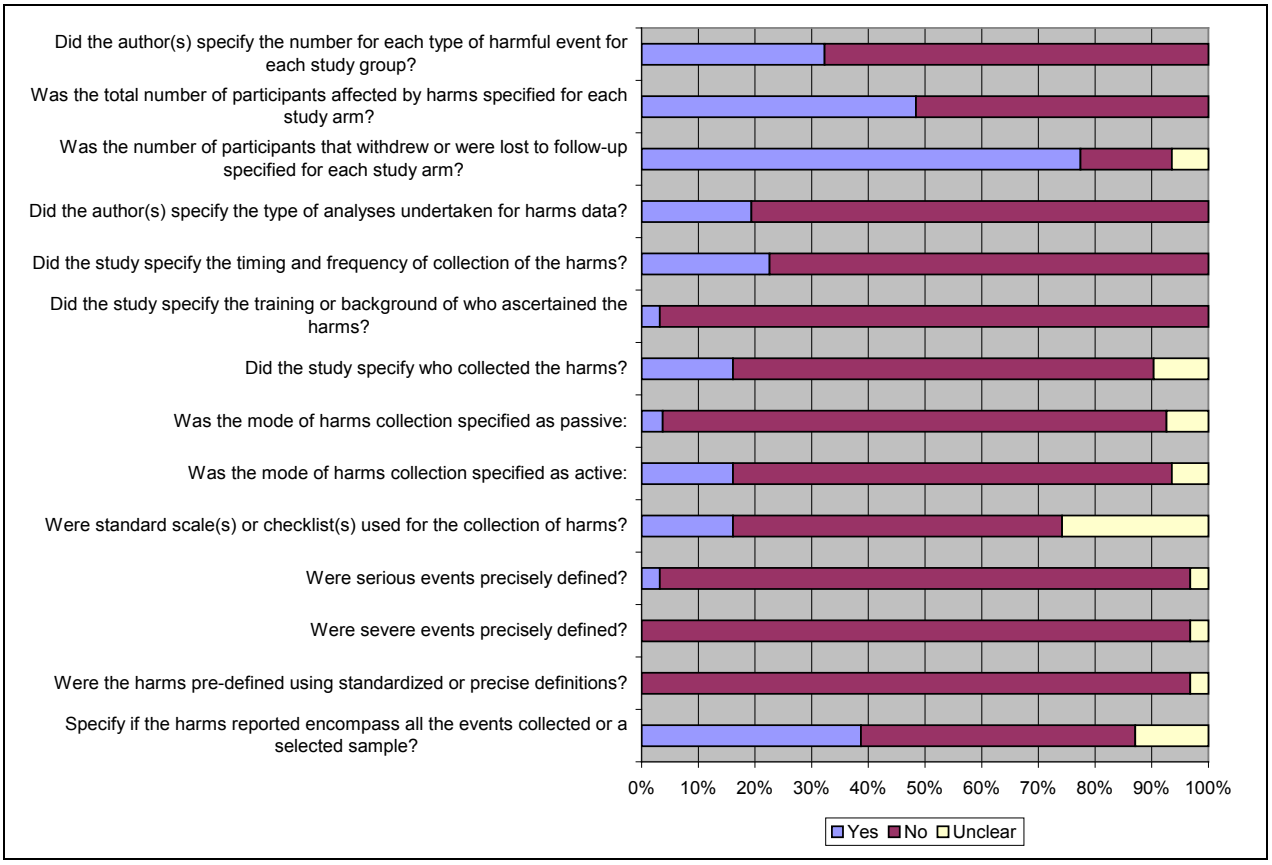
Four studies[108-110,115] evaluated statistical differences in rates of harms, however, two of these evaluated primarily the comparisons for the monotherapy group relative to the combined therapy group.[108,110] Another study[115] evaluated differences between two methods of switching from an SSRI to duloxetine; no statistical differences were found between the two methods.

Monotherapy versus Combined Therapies in Adult MDD

One study[91] reported harms when evaluating monotherapies relative to combined therapies. Only one study[123] was designed to assess the effect of therapies for both efficacy and harms in patients who had excess sleepiness and fatigue (despite previous adequate SSRI treatment); the subjects in this trial were partial responders for the current episode. This study included specific measures of sleepiness and fatigue as part of the primary outcomes.

The method of assessing adverse effects differed greatly among studies with a limited number of studies using standardized methods or the use of scales to assess harms. Figure 3.13 shows the ratings on the McHarm scale for evaluating risk of bias and reporting within comparative studies. Eleven studies (40 percent) indicated that the harms reported were those that were observed in two to three percent,[76] five percent[77-80,110,119,123] or 10 percent of subjects;[108,130,136] however, three of these studies did not report harms specific to the SSRI subgroup.[76-80] The remaining studies did not specify why the harms reported were included or were unclear (20 percent). All but one study[126]provided a priori definitions for serious harms. Similarly, definitions for predefining the harms or how these would be classified as severe were not detailed in any study (Figure 3.14). The mode of collecting harms was unclear or not identified in all but three studies.[120-122,129,133] Who collected reports of harms or their training was rarely specified. Generally, the number of subjects who withdrew were specified per treatment arm, and the total number of adverse events were generally reported.

Figure 3.13 percent of studies evaluated using the criteria for risk of bias using the McHarm scale



Fifteen from 29 studies indicated some type of statistical comparison between groups had been undertaken; however, only five studies[49,109,119,130,133] specified the type of analyses and the remaining ones did not.[75-80,108,110,120,123,125,127] One study[108] showed that weight gain, dry mouth, somnolence, peripheral edema and hypersomnia differed between the combined fluoxetine and olanzapine group relative to the fluoxetine group; rates were higher in the combined group. In this same study no differences in rates of adverse events were shown between the combined group relative to olanzapine monotherapy. Another study evaluating olanzapine showed differences relative to baseline but not between treatment groups.[109]

Another study[110] evaluated differences between two montherapy doses or sertraline relative to sertraline combined with mianserin; statistical differences were shown only for the adverse effect of sedation, with rates being higher in the combined therapy group. One study[123] showed statistical differences in nausea and feeling jittery for the combined SSRI and modofinil group.

There were three studies[75-80] that provided stratified outcomes of benefit for the SSRI subgroup alone. However, these three studies did not provide stratified event rates for harms; as such, the rates of harms are not detailed as they reflect mixed antidepressant effect. For two studies[77-79] the pooled analyses publication[80] indicated that there were no differences between groups due to antidepressant; this pooled analysis found that the combined therapy group with ariprazole had approximately twice the incidence of adverse effects (akathisia, restlessness, insomnia, fatigue,

blurred vision, and constipation). The harms in another study[75] were evaluated statistically and did not differ between antidepressants alone or combined with risperidone groups. Another study found rates of events to be similar between antidepressants versus antidepressants combined with risperidone but differences were not evaluated statistically.

Other adverse events not reported in Table 18 include dry mouth,[108,110,111,115,119,123,124,136] dizziness,[77-80,111,118,123,136] and fatigue,[77-80,108,111,131,132] increased appetite was reported in two studies,[108,119] and cardiovascular problems (hypotension, tachycardia, or bradycardia) was identified in five studies.[118,123,126,127,131,132] For non-pharmacological therapies, most studies assumed that there were no adverse events to report with exercise,[81] cognitive behavioral therapy,[74] or dialectical therapy.[117]

Combined Therapies versus Combined Therapies in Adult MDD

From the six studies comparing combined therapies, none were designed to assess the effect of therapies on harms. The method of assessing adverse effects differed greatly among studies with a limited number of studies using standardized methods or the use of scales to assess harms. Figure 3.14 shows the ratings on the McHarm scale for evaluating risk of bias specific to harms. A single study from six specified that the harms reported represented those that were present in at least 10 percent of subjects.[108] The remaining studies did not specify why the harms reported were included or were unclear (85 percent). No study predefined the harms, or severe or serious harms. The mode of collecting harms, who collected the harms reports, or their training was generally not specified. Generally, the number of subjects who withdrew were specified per treatment arm, and the total number of adverse events were reported.

Figure 3.14 shows the percent of studies evaluated using the criteria for risk of bias using the McHarm scale for combined therapies alone

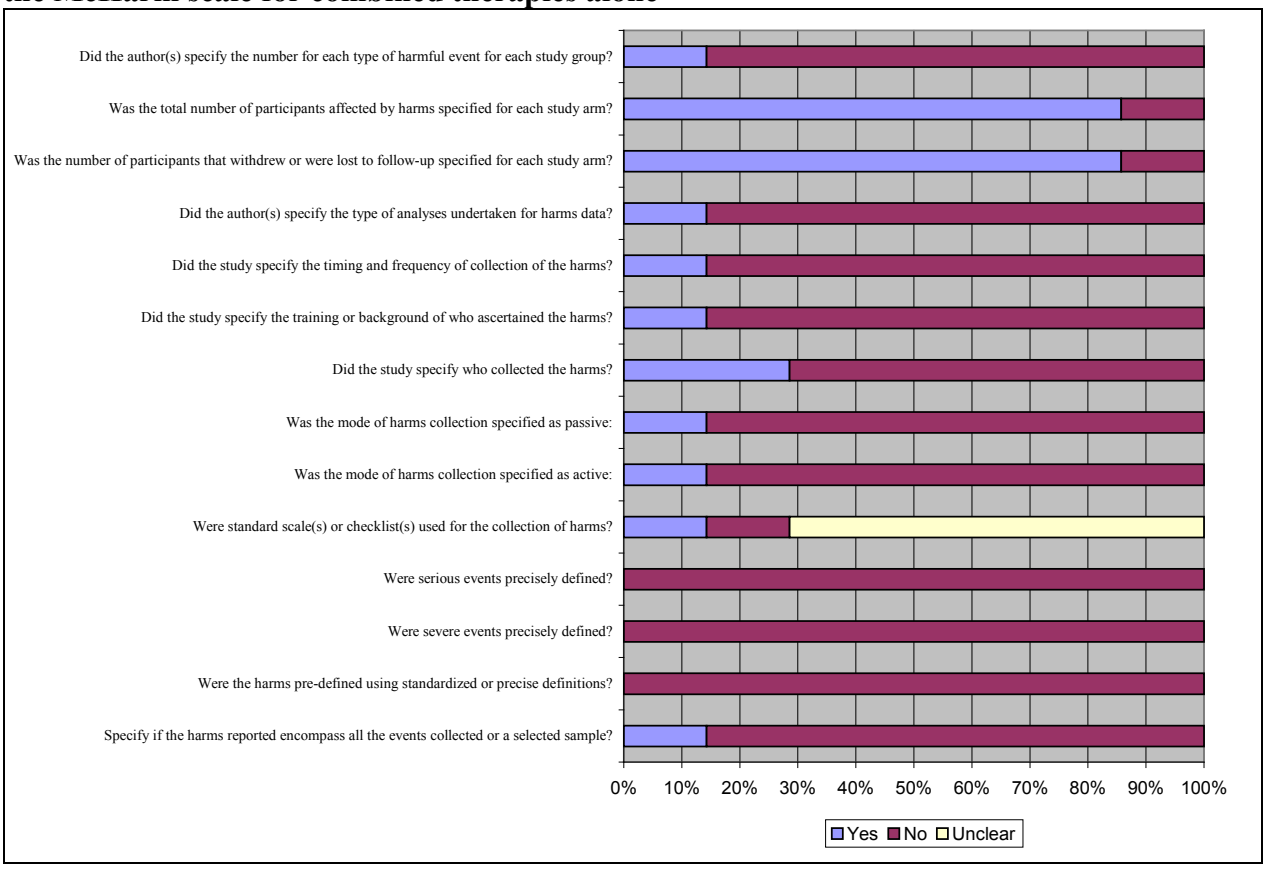


Table 18 shows the rates of reported harms as a function of the treatment arm.The STAR*D cohort reported only the frequency of events and did not specify the type of events or serious events. Two studies explicitly identified that no serious events had occurred,[49,113,116] or that suicide events had been explicitly not occurred.[127] Rates of discontinuation due to adverse events were variable.

A single study[116] reported evaluating statistical differences between groups. Other adverse events not reported in Table 18 include dry mouth,[124,136] dizziness,[136] and fatigue,[131,132] and cardiovascular problems (hyper and hypotension, tachycardia, or bradycardia) was identified in one study.[131,132]

Description of Harms in studies with Dysthymia and Subsyndromal Depression

One study[101] evaluated patients with dysthymia and found no differences in treatment groups (paroxetine versus paroxetine + amisulpride). The presence of galactorrhoea and menstrual disorders were noted in 18 and nine percent of female patients, respectively. These adverse effects were not observed in paroxetine alone group. Other harms reported included low rates of gastrointestinal problems, sexual dysfunction, dry mouth and headache, and some sexual dysfunction. Consistent with studies already described, this study did not predefine harms, serious or severe and indicated that harms were assessed through “spontaneous” notification (passive methods). Nor was the training of the person collecting harms specified or the frequency

and timing of collection. This study did account for all study withdrawals and adequately reported total number of adverse events and as a function of groups for each type of harm.

The single study[100] evaluating harms in patients with subsyndromal depression (following an acute episode) assessed primarily safety and not efficacy. In addition, the study evaluated the relationship between adverse events and the corresponding metabolic status of the isoenzyme CYP 2D6; the rationale for this is that paroxetine is a potent inhibitor of this enzyme which may lead to increased adverse reactions. Adverse effects were measured using the Udvalg for Kliniske Undersøgelser Side Effect Rating Scale (UKU) and the ASEX. The study showed no statistical difference in the UKU scale and the ASEX scale showed an improvement from the first week of treatment in the mirtazapine group. Two subjects from the mirtazapine group discontinued due to problems with insomnia; no drop outs were reported for the paroxetine group.

Description of Harms in studies with Adolescents

The single study that evaluated adolescents who failed to respond to previous treatment with an SSRI found no statistical differences between treatments with regards to the frequency of events, any serious adverse events (including suicide related symptoms), or drop out related to adverse events.[17,95,96] Sleeping difficulties was the only psychiatric adverse event that occurred in greater than five percent of the subjects. Some harms showed a tendency for increased rates with the use of venlaxafine and these included skin rash and cardiovascular events;[17] self injury was also higher in those with higher suicidal ideation.[96] Further analysis of suicidal adverse events showed that predictors of suicidal adverse events were linked with poor response to treatment.

The harms in this study were collected using standardized instruments (4 item Kiddie Schedule of Affective Disorders and the Side Effects form for Children and Adolescents) and collected in an active manner. Reports of serious effects or worsening symptoms were reviewed weekly with the investigative team. Once any concerns for safety were raised, participants were monitored weekly. All subjects completed the standardized safety scales at each pharmacological visit. The reporting of harms was clear, but severe harms were not defined a priori. Withdrawals were well described.

Table 18. Summary of reported rates of harms for studies comparing monotherapy treatments

Study	Duration (weeks)	N*	Comparison and dose (mg/d)	Anxiety N (%)	Sedation N (%)	GI problems N (%)	Sleep problems N (%)	Weight gain N (%)	Head-ache (N (%)	Sexual dysfunc-tion N (%)	With-drawals due to A/E N (%)	Serious events N (%)	Suicide N (%)
ADD SSRI													
Licht[110] 2002	6	99	SER 100 +PBO	NR	12 (12.2)	16 (16.4)	4 (4.1)	3 (3.1)	2 (2)	NR	45 (45)	NR	NR
		98	SER 200 +PBO	NR	16 (16.3)	16 (16.3)	10 (10.2)	2 (2)	7 (7.1)	NR	54 (55)	NR	NR
ADD Non SSRI													
Bondolfi[91] 2006	4	19	PARO 40mg/d	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
		9	VEN 150mg/d	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
Rush[49] 2006	12	238	SER 50-200mg/d	NR	NR	NR	NR	NR	NR	NR	NR	11 (4.6)	2 (0.84)
		250	VEN 37.5-375mg/d	NR	NR	NR	NR	NR	NR	NR	NR	5 (2)	0
		239	BUP 150-400mg/d	NR	NR	NR	NR	NR	NR	NR	NR	6 (2.5)	2 (0.83)
Perahia[115] 2008	10	183	direct switch duloxetine 60-120mg/d	NR	8 (4.3)	33 (18)	13 (7.1)	2 (2.1)	24 (13.1)	6 (3.2)	100 (54.6)	5 (2.7)	2 (2.1)
		185	start-taper switch duloxetine 60-120mg/d	NR	13 (7)	37 (20)	15 (8.2)	6 (3.2)	18 (9.7)	2 (2.1)	93 (50.3)	2 (2.1)	0
ADD Augmenting Agent													
Thase[108] 2007	8	203	FLX 50mg/d	NR	5.3		2.4	6.8	19.4	NR	NR	0	NR
		197	OLZ 6-18mg/d	NR	12.1		11.1	39.7	13.1	NR	NR	0	NR
Shelton[109] 2001	8	8	FLX 20-60mg/d	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
		8	OLZ 5-20mg/d	NR	NR	NR	NR	NR	NR	NR	1 (12.5)	NR	NR
Ferreri[111] 2001	6	38	FLX 20mg/d	NR	0	0	NR	0	3 (7.8)	NR	8 (21)	NR	NR
		34	MIN 60mg/d	NR	5 (14.7)	3 (8.3)	NR	2 (5.8)	2 (5.8)	NR	0	NR	NR
ADD Non-Pharmocological													

Table 18. Summary of reported rates of harms for studies comparing monotherapy treatments

Study	Duration (weeks)	N*	Comparison and dose (mg/d)	Anxiety N (%)	Sedation N (%)	GI problems N (%)	Sleep problems N (%)	Weight gain N (%)	Head-ache (N (%)	Sexual dysfunc-tion N (%)	With-drawals due to A/E N (%)	Serious events N (%)	Suicide N (%)
Trivedi[116] 2006 Thase[113] 2007	12	238	SER 50-200mg/d	NR	NR	NR	NR	NR	NR	NR	NR	11 (4.6)	2 (0.84)
		250	VEN 37.5-375mg/d	NR	NR	NR	NR	NR	NR	NR	NR	5 (2)	0
		239	BUP 150-400mg/d	NR	NR	NR	NR	NR	NR	NR	NR	6 (2.5)	2 (0.83)
		86	Monotherapy Medications	NR	NR	NR	NR	NR	NR	NR	23* (27)	2^	NR
		36	CBT	NR	NR	NR	NR	NR	NR	NR	6 (17)	0^	NR

Abbreviations: PBO = Placebo, SSRI: CIT = Citalopram, SER = Sertraline,dPAX = Paroxetine, FLX = Fluoxetine, SSRI = SSRI, SNRI: AMI = Amitryptiline, BUP = Bupropion, CM = Clomipramine, DLX = Duloxetine, DES = Desipramine, MIL = Milnacipram, MIR = Mirtazapine, VEN = Venlafaxine, Augmenting agents: Anti-CRT = Antiglucicorticoids, AM = Atomoxetine, ARI = Aripiprazole, AAP = Atypical Antipsychotics, BNP = Bezodiazepines, BUS = Buspirone, CP = CP 101 106, ES = Estrogen, FOL = Folate, IN = Inositol, LTG = Lamotrigine, Li = Lithium, ME = Mecamyline Hydrochloride, MIN = Mianserin, MOD = Modafinil, OLZ = Olanzapine, PHN = Phenytoin, PI = Pindolol, QTP = Quetiapine, RIS = Risperidone, STIM = Stimulants, TE = Testosterone, T3 = Tri-iodothyronine, TRP = Tryptophan, ZI = Ziprasidone, Non-Pharmacological: ACUP = Acupuncture, ABNS = Ablative Neurosurgery, BAC = Behavioural Activation Therapy, CBT = Cognitive Behavioural Therapy, CBRSG = CBRSG, DBT = DBT, EX = Exercise, IPT = Interpersoanl Therapy, LT = Light Therapy, MBCT = MBCT, MME = Mono Medications, PSY = Psychotherapy, AME = Augment Medications

KQ3: How do these therapies compare in different populations (e.g., different depressive diagnoses, disease severity, ages, gender, racial and socioeconomic group, and medical or psychiatric co-morbidities)? These subgroups will be considered with respect to the different interventions.

KEY MESSAGES
Overall, there is small number of studies that have evaluated the impact of disease type, disease severity, previous co-morbidities, age, gender, and race on treatment outcomes.
There is some evidence from the STAR*D level 2 cohort that would suggest that persons with concurrent anxiety symptoms have less likelihood of achieving remission.
There is some evidence from the TORDIA trial that milder depression, less family conflict, and absence of suicidal behavior are associated with greater likelihood of a positive treatment response at 12 weeks in adolescents.

Given that there was one study each for patients with dysthymia and subsyndromal depression, this review is limited in meaningfully comparing conclusions across populations with different depressive disorders. There are seven studies that undertook stratified or subgroup analyses evaluating factors that may impact treatment outcomes in adults,[103-105,107,108,110,120-125,144] and one for adolescents.[17,95,96]

Factors affecting treatment response in Adults

Baseline Disease severity. Six studies evaluated the impact of disease severity on treatment outcomes in adults. One study undertook a subgroup analysis on subjects with baseline HAM-D-17 score greater than 17 and found that the group with combined treatment (SSRI + modafinil) had statistically significant greater reduction ($p = 0.05$) relative to the SSRI group alone. Another study[125] found that subjects with an initially higher MADRS score tended to show greater reductions in MADRS overall ($p = 0.04$), or within the first 2 weeks of treatment (MADRS (>30) in the combined therapy group (fluoxetine/ citalopram + buspirone) relative to subjects in the SSRI group with higher initial scores. One study[106,107,124] found that lower baseline HAM-D-17 score was predictive of response for the fluoxetine group ($p = 0.008$) and the lithium augmentation group ($p = 0.04$) but not the desipramine group; a re-analysis found that an OR = 0.85 (95% CI 0.76 to 0.96) was found for any augmentation strategy relative to a dose increase in fluoxetine.[106] One of these studies found that age of onset of depression was predictive of response ($p = 0.009$).[106,107,124]

Analysis of level 2 STAR*D cohort found that subjects with severe depression (QID-SR 16 or greater) were less likely to achieve remission (OR = 0.34 (0.22 – 0.52); however, this aspect was not valuable in assisting clinicians in recommending any monotherapy treatment (sertraline, venlaxafine, buproprion).[103] Greater baseline symptom severity was also associated with greater rates of attrition.[105]

Two studies evaluated baseline HAM-D scores (>23)[110] and baseline severity[120-122] and showed that these did not affect treatment response.

Previous history of failure. Two studies evaluated previous history of failure. One study undertook a subgroup analysis evaluating the drug class of previous failure (SSRI versus other); this study showed differences with the combined olanzapine-fluoxetine group achieving a statistically significant greater reduction on the MADRS relative to the fluoxetine or olanzapine monotherapy groups.[108] This trend was observed in the non-SSRI group for those with at least one previous failure, but only for olanzapine and not fluoxetine.[108]

In the STAR*D level 2 cohort, intolerance to citalopram (OR = 1.57 (1.11-2.21)) or response to citalopram during step 1 (OR = 2.78 (1.77-4.38)) increased the likelihood of remission; however, this was not practically helpful to clinicians in selecting one monotherapy treatment over the other.

Comorbidities. The STAR*D cohort analysis for level 2 subjects on monotherapies (sertraline, venlaxafine, buproprion), showed that remission was less likely in patients with other concurrent psychiatric disorders (specifically, panic or posttraumatic disorders, generalized anxiety disorder, obsessive compulsive disorders, obsessive-compulsive disorder, social phobia, anxious or melancholic features).[103] The overall OR for presence of anxious, atypical, or melancholic features were 0.30 (0.20 to 045), 1.04 (0.67-1.61) and 0.43 (0.25-0.73) respectively.[103]

A more detailed analysis of the STAR*D level 2 cohort showed that rates of remission were significantly less for anxious patients relative to non-anxious patients across all five pharmacological treatment arms (both monotherapy and combined therapy).[104] Logistic regressions however, indicated only a moderate effect of anxiety, suggesting that there was no advantage of one treatment over another in subjects with anxious depression.[104]

One study showed no significant difference on treatment response for subjects with melancholic features.[110]

Age. Two studies showed no statistical difference when the impact of age on treatment response was evaluated.[107,110,124] Analysis of the STAR*D level 2 cohort showed that age younger than 35 increased the likelihood of remission (OR= 1.43-1.81).[103] In contrast, younger age was associated with attrition for the augmentation treatment group.[105]

Gender. Three studies evaluated gender [107,110,120-122,124] and showed no statistical difference on treatment response. The STAR*D cohort at level 2 estimated an OR = 0.96 (0.69-1.35); overall gender was not an important factor in helping to select the optimal monotherapy.[103]

Race. Non-white races were associated with greater rates of attrition for level 2 STAR*D subjects;[105] conversely, white race was associated with greater likelihood of remission.[103]

Factors affecting treatment response in Adolescents

An analysis of the TORDIA trial[17,95,96] for predictors of treatment response showed that milder depression, less family conflict, and absence of suicidal behavior were associated with greater likelihood of a positive treatment response at 12 weeks. In the context of combined treatment of CBT with antidepressant, adolescents with no history of abuse and few co-morbidities, had a greater probability of a positive response.[95] Older youths (age 18-19) (OR 3.7 (95% CI 1.2-12.0)) with more comorbidities are more likely to benefit from combined treatment.[95]

KQ4: What is the range of recommended clinical actions following the failure of one adequate course of SSRI based on current (<5 years) clinical practice guidelines?

KEY MESSAGES
There were 23 Clinical Practice Guidelines (CPG) (14 for adults, seven for adolescents and two including both) providing recommendations for patients with MDD. Four CPG for adults and three for adolescents did not provide any recommendations for patients with previous inadequate responses. Four guidelines included patients with dysthymia and subsyndromal depression but no recommendations for these subgroups who had failed previous treatment for both adults and adolescents. The majority of CPG did not specify a definition for inadequate response.
All CPG for adults and adolescents were applicable to patients from primary care and outpatient settings; a smaller number indicated applicability to inpatient settings.
For adults, the majority of CPG did not specify any type of antidepressant when recommending switching to monotherapy strategies. Increasing the dose and duration was frequently recommended but the interval or change in dose was not specified. When combined therapy was recommended there was a greater tendency to specify the drug for adding antidepressants. However, there was great variability in the augmenting agents recommended.
For adolescents, there was approximately equal number of CPG that specified which agents to consider for monotherapy and which to consider for combined therapies. There was a preference in commence treatment using non-pharmacological treatments. Some guidelines cited adult evidence as the evidentiary basis for suggesting treatment strategies.

There were a total of 23 clinical practice guidelines (CPG) sponsored by unique organizations and described in 30 publications.[18,37,57-59,145-169] Note that CPGs can be published as a comprehensive single document with numerous recommendations for different interventions or CPG can be published as multiple documents related to different interventions but sponsored by the same organization and published the same year. For the purposes of this review, we consider each sponsoring organization as the unit of analysis for a CPG, as the methodology and recommendation format is the same across related multiple publications. Three publications[162-164] do not contain recommendations, but are eligible as they provide supporting documentation for

those publications that provide treatment specifications from the Canadian Network for Mood and Anxiety Treatments (CANMAT) guideline.[57,58,161] One publication is a summary companion paper[170] of another CPG from the American College of Physicians (ACP).[59] Two publications[145,146] are from the Guidelines for Adolescent Depression in Primary Care (GLAD-PC). Figure 2 shows that 56 guidelines were excluded because of the following: 1) publication prior to 2005 (N = 46); 2) exclusive focus on diagnosis or screening rather than treatment (N = 5); and 3) not a population of interest (N = 5).

There were seven CPG that were specific only to adolescents,[18,37,146-150,154] 14 CPG for adults alone,[57-59,151,153,154,156-161,165-168] and two applicable to both adults and adolescents.[155,169]

CPG specific for MDD, Dysthymia, and Subsyndromal Depression in Adults

Characteristics of CPG for Adults. Table 19 shows the characteristics of the CPG as a function of country of origin, setting, and intended users. All 16 CPG were applicable to adults with MDD. Two CPG[153,155] had some recommendations applicable to patients with Dysthymia and one also specified treatment for Subsyndromal depression;[155] however, none of the recommendations within these two CPG were specific for dysthymia or subsyndromal depression patients who had failed to respond to previous treatment. One CPG summarizes evidence on pharmacological treatment that includes both MDD and Dysthymia but presents no recommendations specific to Dysthymia.[59] One CPG considered sub threshold persistent symptoms as a distinct subgroup of patients;[167] the recommendations did not differ from those with inadequate response in patients with mild to moderate MDD.

All but two of the 16 guidelines considered a variety of treatment interventions; these two CPGs evaluated solely pharmacological interventions,[59] and computerized CBT.[168] The other CPG's included treatment recommendations that provided treatment strategies including a variety of pharmacological, psychological, and complementary and alternative (CAM) interventions; however the majority of recommendations were not applicable to patients who have had inadequate responses to previous pharmacological treatment. When recommendations were specific to patients who had previous inadequate response, none were distinguished by those that had failed to different classes of anti-depressants.

All CPG for adults were applicable to patients from primary care and outpatient settings; six guidelines indicating applicability to inpatient settings (Table 19). All CPG were intended primarily for or applicable to primary care practitioners, with the exception of one CPG that was developed specifically for psychiatrists.[159] The majority of guidelines were undertaken in the US (N = 6), one was sponsored in Singapore,[166] and one[157] by the World Federation of Biological Psychiatry.

Table 19. Characteristics of included CPG showing country, disorder type included, setting, and intended guideline users

	Adult United States	Adult Canada	Adult United Kingdom	Adult New Zealand/ Australia	Adult Other

Disorder					
MDD	Jaehne[151] Qaseem[59] Karasu[153] Steinman[156] Davidson[158]	Ravindran[58] Parikh[57] Lam[161] Conn[165]	Anderson[155] NIHCE[167] NIHCE[168]	Ellis[159] Malhi[160] NZGG[169]	Bauer[157] Mahendran[166]
Dysthymia	Karasu[153] [Qaseem[59]]*		Anderson[155]		
Subsyndromal			Anderson[155]		
Setting					
Primary Care	Jaehne[151] Qaseem[59] Horsley[152] Karasu[153] USPSTF[171] Steinman[156] Davidson[158]	Ravindran[58] Parikh[57] Lam[161] Conn[165]	Anderson[155] NIHCE[167] NIHCE[168]	Ellis[159] Malhi[160]	Bauer[157] Mahendran[166]
Outpatient MH	Jaehne[151] Qaseem[59] Karasu[153] Anderson[155]	Parikh[57] Conn[165]	Anderson[155] NIHCE[168]	Ellis[159] Malhi[160]	
Inpatient MH	Qaseem[59] Karasu[153]	Parikh[57] Conn[165]	Anderson[155]	Ellis[159]	
Other					
Intended Users					
Primary Care Physicians	Jaehne[151] Qaseem[59] Karasu[153] USPSTF[171] Steinman[156] Davidson[158]	Ravindran[58] Parikh[57] Lam[161] Conn[165]	Anderson[155] NIHCE[167] NIHCE[168]	Malhi[160]	Bauer[157] Mahendran[166]
Mental Health Specialists	Jaehne[151] Qaseem[59] Karasu[153] USPSTF[171] Steinman[156] Davidson[158]	Parikh[57] Lam[161] Conn[165]	Anderson[155] NIHCE[167]	Ellis[159] Malhi[160]	Mahendran[166]
Allied Mental Health disciplines	Jaehne[151]	Parikh[57]			

* Dysthymia population included in the CPG but no recommendations are specific to Dysthymia but only for MDD.
Abbreviations: NIHCE = National Institute for Health and Clinical Excellence, NZGG = New Zealand Guidelines Group, USPSTF = United States Preventive Task Force Services

Inadequate Response. From 16 CPG, 11 did not report any specific definitions for defining adequate response or remission within the guideline. Four CPG did not include recommendations specific to failed response populations and as such, a definition may not have been necessary.[156,158,165,168] The remaining seven CPG did not report a specific definition and as such inadequate response is left open to variable interpretation.[59,154,155,157,159,160,167]

Seven CPG defined response as a 50 percent or greater reduction in symptoms (as measured on a standardized rating scale) and partial response is defined as a 25 to 50 percent reduction in symptoms;[57,58,151,157,161,164] One CPG specified that the measure should be a change in the Patient

Health Questionnaire – 9.[169] The CANMAT CPG recommendations were intermingled with order of treatment and lack of adequate response. First line treatment is identified as those interventions for which there is the best evidence of efficacy balanced with good safety and tolerability. Second and third line treatments are defined as those reserved for situations where first line treatments are not indicated, cannot be used, or when first line treatments are not effective. As such, for the CANMAT guidelines specific to CAM[58] and psychological therapies[161] there failed to respond populations are not identified clearly within the body of the recommendations; we must assume that second and third line therapies are applicable to those that failed previous pharmacological treatments.

For those CPG's that did report a formal definition of inadequate response, only two CPG provided clear indications for differential treatment strategies for those with partial versus non-response.[151,153] Seven CPG indicated that the definition of inadequate response was linked to failure following time intervals varying from 2 to 4 weeks,[155,157,166], 4 to 6 weeks of significant improvement,[169] ,4 to 8 weeks,[153] and 6 to 8 weeks of partial improvement.[59,166]

Quality Assessment of CPG for Adults. Table 20 shows the domain scores for the AGREE II ratings of the CPG. The AGREE II is based on six domains of methodology for the guideline process. Two guidelines rated poorly overall across all but the domain on "scope and purpose".[158,166] All CPG scored high for scope and purpose (Domain 1) (range 86 to 100 percent). The lowest scores were observed in the "clarity of presentation" (Domain 4) (range 17 to 42 percent). This domain evaluated whether the recommendations were clear and unambiguous, such that options were clearly presented, and key recommendations easily identifiable.

Stakeholder involvement (Domain 2) had wide ranging scores varying from 39 to 92 percent, and only five from 16 CPG indicated that patient's views and preferences had been sought (score five or greater).[151,157,159,167,168] For the domain of "rigor of development" (Domain 3), scores varied from 43 to 85 percent; all but three CPG[151,167,168] did not indicate a process for updating the guideline. When considering the "applicability" domain (Domain 5) scores varied from 13 to 78 percent. The majority of CPG scored poorly for two criteria within this domain: 1) consideration of potential resource implications of applying their recommendations; and 2) presenting monitoring or auditing criteria. For the domain regarding "editorial independence" (Domain 6), scores were highly variable and ranged from four to 96 percent. In particular, the competing interest of the guideline development group were not consistently recorded.

Although the AGREE II evaluates the methodology of the guideline process, it cannot evaluate the scientific merit and overall quality of the recommendations. All of the CPG had methods to establish the strength of the evidence but none could be compared to each other. Most systems of grading the strength of the evidence included aspects of study design (for example, RCT) or number of studies or confidence of treatment; most included a level that reflected consensus or expert opinion for some recommendations.

Table 20. Scores from the AGREE II for CPG for adults

Author	Organization	Scope and Purpose (Domain 1)	Stakeholder Involvement (Domain 2)	Rigor of Development (Domain 3)	Clarity of Presentation (Domain 4)	Applicability (Domain 5)	Editorial Independence (Domain 6)
Jaehne[151] 2009	Institute for Clinical Systems Improvement	97.22	88.89	82.29	29.17	63.33	95.83
Qaseem[59] 2008	American College of Physicians	94.44	58.33	59.38	25.00	28.33	87.50
Karasu[153] 2009	American Psychiatric Association	100.00	61.11	71.88	33.33	26.67	45.83
National Guideline Clearing House[154] 2004	National Guideline Clearing House	97.22	63.89	80.21	37.50	21.67	79.17
Steinman[156] 2007	CDC	97.22	63.89	66.67	41.67	68.33	50.00
Davidson[158] 2006	National Heart, Lung, and Blood Institute	91.67	61.11	42.71	33.33	16.67	12.50
Ravindran[58] 2009	Canadian Network for Mood and Anxiety Treatments	86.11	38.89	68.75	29.17	21.67	70.83
Parikh[57] 2009	Canadian Network for Mood and Anxiety Treatments	94.44	58.33	69.79	16.67	20.00	50.00
Lam[161] 2009	Canadian Network for Mood and Anxiety Treatments	97.22	63.89	85.42	29.17	15.00	70.83
Conn[165] 2006	Canadian Coalition for Seniors™ Mental Health	100.00	58.33	82.29	33.33	25.00	37.50
Anderson[155] 2008	British Association for Psychopharmacology	91.67	58.33	72.92	41.67	20.00	12.50
NICE CBT[172] 2009	National Institute for Clinical Excellence	97.22	88.89	77.08	33.33	70.00	50.00
NICE[168] 2008	National Institute for Clinical Excellence	94.44	83.33	78.13	25.00	78.33	45.83
Ellis[159] 2004	RANZCP	100.00	91.67	82.29	41.67	36.67	66.67
Malhi[160] 2009	NSCCMHDA	100.00	66.67	69.79	41.67	26.67	66.67
New Zealand Guidelines Group[169] 2008	Ministry of Health & New Zealand Guidelines Group	100.00	66.67	69.79	33.33	50.00	100.00
Bauer[157] 2007	World Federation of Societies of Biological Psychiatry	91.67	83.33	81.25	37.50	33.33	12.50
Mahendran[166] 2005	Ministry of Health Singapore	94.44	41.67	0.00	16.67	13.33	4.17

Recommendations of CPG for Adults. Three CPG specific to MDD did not provide any recommendations for adult patients who had failed to respond to treatment. Two of these CPG were specific to elderly patients in the community,[156] and in long term care homes.[165]. One CPG

had recommendations for patients with depression and cardiovascular disease[158] but none for those who had inadequate response.

Table 21 shows the recommended strategies for both monotherapy and combined therapies. Attempts were made to identify any recommendations with regards to specific medications that were highlighted; however, for some guidelines it was not clear if the text following the recommendation (for example, "Switch antidepressants") was a selective summary of the available evidence or actually recommendation for action. The CANMAT CPG recommended a stepped approach to treatment, intending a particular sequence of interventions (for example, second and third line therapies); however, there were several options within each of these categories.[57,58,161] Other CPG's did not explicitly indicate an order of treatment, other than cautioning to optimize initial treatment. Two CPG did not explicitly recommend a change in dose or duration[59,157] Two CPG distinguished between partial vs. non-response and specified different treatment approaches to these.[153,161] In general, when treatment failed subjects or treatment resistance patients were identified, the recommendations did not differ except for recommendations for ECT and vRMS.

Table 21. Recommendations for treatment in CPG that identified strategies for those that failed response (N = 12)

		MONOTHERAPY					COMBINED THERAPY				
	Starting interval (weeks)	**Dose or duration Change**	**Switch Other SSRI**	**Switch Non-SSRI**	**Switch to Augmenting Agent**	**Switch Non-Pharm**	**Add Augmentor**	**Add Other SSRI**	**Add Non-SSRI AD**	**Add Non-Pharm**	**Add Other**
United States											
Jaehne[151] 2009	6	X	X	X		PSY* LT* AC ECT* VNS* DBS*	NS * TCA + T3* TCA + LI* AD + AR* AD + AAP*		SSRI + BU* SSRI + MI* SSRI +TCA*		
Qaseem[59] 2008	6-8		CI. FL, FU, PA, SE	MI			X	X	X	X	
Karasu[153] 2009	4-8	X	X	X		PSY ECT	NS AD + LI or T3, or STIM	AD	AD	PSY ECT	
546	NS	X	X	X		PSY	SSRI + LI (300-600mg/d)		SSRI + DE	PSY	
Canada											
Ravindran[58] 2009						OM3 SAM-e DHEA FA	LT EX Yoga SleepD OM3				AD + CBT or IPT
Parikh[57] 2009	NS					BAC CBASP IPT MBCT					
Lam[161] 2009	1-4	X	ES SE VE	DU MI MIL **** AMT or CM or MAO			AD + AR or LI or OL or RI *** AD + QU or T3 or MI *** AD + BS or MO or ZI or	X	AD + BU or MT or		

		MONOTHERAPY					COMBINED THERAPY				
	Starting interval (weeks)	**Dose or duration Change**	**Switch Other SSRI**	**Switch Non-SSRI**	**Switch to Augmenting Agent**	**Switch Non-Pharm**	**Add Augmentor**	**Add Other SSRI**	**Add Non-SSRI AD**	**Add Non-Pharm**	**Add Other**
							STIM				
United Kingdom											
Anderson[155] 2008	4 - 8	X	X	X		CBT PSY EX, ECT rTMS VNS, ABNS	SSRI + LI or OL or AR, TCA + T3 AD + IA or TR, or MO, STIM, ES or AG or OM3 or FO,	AD + MT			
National Institute for Health and Clinical Excellence[167] 2009	6-8	X	X	X			AD + AR or LI or OL or RI or MI or QU	X	AD + MT		
National Institute for Health and Clinical Excellence[168] 2008											
New Zealand/ Australia											
Ellis[159] 2004	NS	X	AD	AD		CBT	TAC + LI SRI + LI or T3 or PI		SSRI + TCA		
Malhi[160] 2009	2 – 6					CBT ECT	AD + LI or T3, or ATA or BE				
New Zealand Guidelines Group[169] 2008											
Other											

		MONOTHERAPY					COMBINED THERAPY				
	Starting interval (weeks)	**Dose or duration Change**	**Switch Other SSRI**	**Switch Non-SSRI**	**Switch to Augmenting Agent**	**Switch Non-Pharm**	**Add Augmentor**	**Add Other SSRI**	**Add Non-SSRI AD**	**Add Non-Pharm**	**Add Other**
Bauer[157] 2007	2 4 weeks		AD	AD			AD + LI or T3, or ATA	X	X		
Mahendran[166] 2005	4-8	X	X	X			AD + Li or T3				

* Applicable to Partial responders or treatment resistance and may require consultation with a specialist

+ Specific for treatment resistant (definition not specified)

Treatment refractory (definition not specified)

The time interval indicates the number of weeks following the first line therapy attempt to initiate new treatment strategy

CPG specific for MDD and Dysthymia in Adolescents

Characteristics of CPG for Adolescents. There were seven CPG that were specific only to adolescents,[18,37,145-150] and two applicable to both adults and adolescents.[155,169] Table 21 shows the characteristics of the adolescent CPG as a function of country of origin, setting, and intended users.

All nine CPG were applicable to adolescents with MDD. Two CPG had some recommendations applicable to patients with Dysthymia, [37,155] and one also specified treatment for subsyndromal depression[37] in adolescents. However, none of the recommendations were specific to those who had failed previous treatment.

All CPG for adolescents were applicable to patients from primary care and outpatient settings; two guidelines indicating applicability to inpatient settings (Table 22). All CPG were intended primarily for or applicable to primary care practitioners, and three to specialists[37,149,155] and allied mental health workers.[37]

The majority of guidelines were undertaken in the US (N = 6), one in the UK,[155] and two in Australia and New Zealand.[147,169]

Table 22. Characteristics of CPG based on the type of disorder, the setting, and intended users

	Adolescent United States	**Adolescent Canada**	**Adolescent United Kingdom**	**Adolescent New Zealand/ Australia**
Disorder				
MDD	Zuckerbrot[145] Cheung[146] US Preventive Services[18] Birmaher[37] Hughes[148] Gallagher[149]		Anderson[155] National Institute for Clinical Excellence[150]	Dudley[147] New Zealand Guidelines Group[169]
Dysthymia	Birmaher[37] Birmaher[173]		Anderson[155]	
Subsyndromal			Anderson[155]	
Setting				
Primary Care	Zuckerbrot[145] Cheung[146] US Preventive Services[18] Birmaher[37] Hughes[148] Gallagher[149]		Anderson[155] National Institute for Clinical Excellence[150]	Dudley[147] New Zealand Guidelines Group[169]
Outpatient MH	Anderson[155] Birmaher[37] Gallagher[149]			
Inpatient MH	Anderson[155] Birmaher[37]			
Other				

Intended users				
Primary Care Physicians	Zuckerbrot[145] Cheung[146] US Preventive Services[18] Birmaher[37] Hughes[148] Gallagher[149]		Anderson[155] National Institute for Clinical Excellence[150]	Dudley[147] New Zealand Guidelines Group[169]
Mental Health Specialists	Anderson[155] Birmaher[37] Gallagher[149]			
Allied Mental Health disciplines	Birmaher[37]			

Inadequate Response. Only two CPG provided definitions of inadequate response and this was characterized as failure of remission over a period of at least 2 weeks and less than 2 months with no or very few depressive symptoms using a children's global assessment scale/interviews[37] or as failure to have significant level of improvement from 4 to 6 weeks.[169]

Quality Assessment of CPG for Adolescents. Table 23 shows the domain scores for the AGREE II ratings of the CPG. One guideline rated poorly across three domain domains 3 to 5).[149] All CPG for adolescents scored high for "scope and purpose" (Domain 1) (range 89 to 100 percent). The lowest scores were observed in the "clarity of presentation" (Domain 4) (range 0 to 42 percent); this domain evaluated whether the recommendations were clearly presented.

The remaining domains showed highly varying scores from 4 to 97 percent in the stakeholder involvement (Domain 2), and the views of the patients and public were sought in only two CPG[148,150] (score six or greater). For the domain of "rigor of development" (Domain 3), scores varied from 21 to 92 percent; only one CPG[150] indicated a process for updating the guideline. When considering the "applicability" domain (Domain 5) scores varied from 15 to 85 percent; the majority of CPG scored poorly for two criteria within this domain: 1) consideration of potential resource implications of applying their recommendations, and 2) presenting monitoring or auditing criteria. For the domain regarding "editorial independence" (Domain 6), scores were highly variable and ranged from thirteen to 100 percent; in particular, the competing interest of the guideline development group were not consistently recorded.

As expected the CPG for adolescents had varying methods to establish the strength of the evidence and none could be compared to each other. Similar to adult rating systems, most CPG used grading systems that included aspects of study design (for example, RCT) or number of studies or confidence of treatment; most included a level that reflected consensus or expert opinion for some recommendations.

Table 23. The AGREE II ratings for the 6 domains in CPG specific to adolescents

Year	Organization	Scope and Purpose (Domain 1)	Stakeholder Involvement (Domain 2)	Rigor of Development (Domain 3)	Clarity of Presentation (Domain 4)	Applicability (Domain 5)	Editorial Independence (Domain 6)
United States							
Zuckerbrot[145] 2009	American Academy of Pediatrics	91.67	61.11	79.17	33.33	73.33	75.00
Cheung[146] 2007	GLAD-PC	100.00	63.89	81.25	25.00	36.67	50.00
U.S. Preventive Services[18] 2009	U.S. Preventive Services Task Force (USPSTF)	88.89	58.33	34.38	33.33	33.33	20.83
Hughes[148] 2007	Texas Department of State Health Services (DSHS)	100.00	91.67	47.92	41.67	48.33	62.50
Birmaher[37] 2007	American Academy of Child and Adolescent Psychiatry	91.67	55.56	65.63	33.33	18.33	16.67
Gallagher[149] 2005	NR	100.00	41.67	21.88	.00	6.67	66.67
United Kingdom							
Anderson[155] 2008	British Association for Psychopharmacology	91.67	58.33	72.92	41.67	20.00	12.50
National Institute for Clinical Excellence[150] 2005	National Institute for Clinical Excellence	100.00	97.22	92.71	37.50	81.67	91.67
Australia New Zealand							
Dudley[147] 2008	NHMRC	91.67	55.56	40.63	33.33	15.00	12.50
New Zealand Guidelines Group[169] 2008	Ministry of Health & New Zealand Guidelines Group	100.00	66.67	69.79	33.33	50.00	100.00

Recommendations of CPG for Adolescents. Three from nine CPG for adolescents did not provide any specific recommendations for adolescents who had failed to respond to previous treatment. One component of a CPG from the GLAD-PC focused only on identification and initial management.[145] One CPG focused only on psychotherapy interventions and did not provide any recommendations specific to those who fail previous treatment.[149] Another CPG from the United States Preventative Services Task Force (USPSTF) focused primarily on recommendations for screening and initial management.[18] One guideline indicates that there is lack of evidence for the management of next steps of treatment for adolescents and provide no further indications.[155]

Two CPG provided recommendations following failure of psychological interventions. One CPG[169] that evaluated treatment for MDD in both adult and adolescent, directed primary care practitioners to refer to secondary mental health services following lack of substantial

improvement after six to eight weeks of supportive and psychological therapies; similarly the recommendation was to seek adolescent psychiatric consultation if the use of an anti-depressant was desired. Two CPG on adolescents[37,150] provided recommendations for patients who had failed to respond to psychotherapy or had more complicated depressions; failure to pharmacological treatment was not clarified for mild depression and recommended only for moderate to severe MDD.

Table 24 shows the proposed treatment options for adolescents with MDD. Three CPG[37,147,148] note the lack of evidence for adolescents but cite adult evidence as the rationale for treatment strategy for switching and augmentation strategies in particular. Once CPG makes clear recommendations to avoid the use of paroxetine and venlaxafine in adolescents 12 to 18.[150]

Table 24. Recommendations for treatment in CPG that identified strategies for those that failed response (N = 5) in adolescents

		MONOTHERAPY					COMBINED THERAPY				
	Starting interval (weeks)	**Dose or duration Change**	**Switch Other SSRI**	**Switch Non-SSRI**	**Switch to Augmenting Agent**	**Switch Non-Pharm**	**Add Augmentor**	**Add Other SSRI**	**Add Non-SSRI AD**	**Add Non-Pharm**	**Add Other**
USA											
Cheung[146] 2007	6-8	X	X	X^		PSY					CON
Hughes[148] 2007	NS	X	CIT, ESC, PAR*, FLX, SER	BUP, VEN. MRT, DUL		ECT	SSRI + Li		SSRI + BUP, MRT		
Birmaer[37] 2007	2 – 8	X	X	X		CBT or IPT	AD + Li or T3			AD + CBT or IPT	
UK											
National Institute for Clinical Excellence[150] 2005			SER, CIT	X^						SSRI + PSY	CON
New Zealand/ Australia											
Dudley[147] 2008	4	FLU	SER, CIT			CBT ECT	SSRI + Li or T3				CON

* PAR not recommended for preadolescents
Must have failed two SSRI and augmentation precedes switch to non-SSRI
ECT if pharmacological treatment fails or depression is severe
Considered only for severe or psychotic cases
CON = Consultation with mental health specialist
X^ = not ideally recommended but can be considered
The time interval indicates the number of weeks following the first line therapy attempt to initiate new treatment strategy

Chapter 4. Discussion

Overview

Pharmacological agents are one of several treatment modalities used to treat major depressive disorder. One of the most frequently utilized classes of antidepressant medications is the selective serotonin reuptake inhibitors (SSRIs). The rate of treatment response following first-line treatment with SSRIs is moderate, however, varying from 40 to 60 percent; remission rates vary from 30 to 45 percent.[19] This comparative effectiveness review has summarized the evidence for management of patients subsequent to a trial of an SSRI that did not result in an adequate response.

KQ1. Among adults and adolescents with major depressive disorder (MDD), dysthymia, and subsyndromal depression, who are started on an SSRI and who are compliant with treatment but fail to improve either fully, partially, or have no response, what is the benefit (efficacy or effectiveness) of monotherapy and combined therapy?
KQ1-A. How does the efficacy/effectiveness vary between the different monotherapies and combined therapies?

Adults

As noted previously, the data comparing single treatments against each other following inadequate response to SSRI are limited; many relatively straightforward clinical questions remain to be addressed. When patients are being treated with an antidepressant and not improving, the first step is often to 'optimize' the treatment, sometimes defined as an adequate dose for an adequate duration of time; there is, however, no consensus on exactly how long a patient should be treated with a medication before there is a decision made that the response has been inadequate.

Even the issue of determining what defines an inadequate trial of an SSRI is therefore not straightforward; while most studies used adequate doses of medications (as defined when the medication receives an indication from a federal authority) the duration of treatment with the SSRI before a judgment is made regarding the inadequacy of the response, was highly variable across the trials that were reviewed. The duration of treatment with an SSRI prior to the determination of an inadequate response ranged from a mean of approximately 4 weeks to a mean of approximately 12 weeks. Although adequate doses may be those defined in product monographs, there has been uncertainty regarding the maximal dose to which many common antidepressants should be prescribed. Despite this, a survey conducted a decade ago suggested that the preferred intervention of clinicians following inadequate response to SSRI was a dose

increase. Few studies have examined whether an increase in dose is associated with a comparable probability of response or remission as selecting another strategy, such as a medication switch or an add-on therapy.

Once a decision is made to move beyond optimizing the SSRI, clinicians have several options available, including switching to a new medication (either of the same antidepressant class or a different class), adding a second antidepressant or adding another agent that itself might not be recognized as an antidepressant in monotherapy. In recent years, the line between 'augmentation' agents and 'antidepressant' medications has been questioned, leading some investigators to suggest that the standard terminologies of 'combination' therapy to refer to multiple antidepressant treatments being used at once (either pharmacological or other) and 'augmentation' therapy to refer to an antidepressant used in conjunction with another non-antidepressant therapy should be collapsed and called 'add-on' therapy. Regardless of preferences for nomenclature, this review has highlighted that there is an extremely limited evidence base to support clinical decision making around any of these strategies.

A common treatment approach following inadequate response to first treatment is to switch medications. Despite STAR*D, there remains limited evidence to determine whether a patient who elects to have a medication switch as the preferred treatment strategy following an inadequate response to one SSRI can be switched to another SSRI with equal likelihood of response compared to switch to a medication from another class. The STAR*D trial suggested that this might be the case, at least when comparing sertraline, venlafaxine and buproprion, but given the frequency with which this question arises in clinical practice, a more substantive evidentiary base on which to make this decision appears warranted.

Another common clinical issue following inadequate treatment response is whether to add a medication, either another antidepressant or a non-antidepressant agent, traditionally called an augmenting agent. Adding a second antidepressant to an SSRI is not uncommon in clinical practice, particularly if patients have had a partial response (at least 20% improvement). Altamura and colleagues[131,132] have compared intravenous citalopram to intravenous clomipramine following inadequate response to SSRIs, but these trials were preliminary and the short term use of intravenous medication does not address the more typical situation in which patients have a second oral medication added to the SSRI.

Traditional augmentation strategies comprised the bulk of studies meeting criteria for inclusion in this review. Most trials investigated whether adding a new agent to ongoing SSRI treatment was preferable to adding a placebo to ongoing SSRI treatment. Therefore, in most instances what was compared was monotherapy with the initial SSRI against co-therapy with the original SSRI and a new agent. Although the majority of studies fell into this category, there are a limited number of studies for any particular augmenting agent, limiting the strength of the results. The array of agents studied meant that it was difficult to make informed decisions regarding specific classes of medications. Additionally, there are very few studies that examined augmentation compared to switching strategies, which is a clinically relevant question. That is, many clinicians would likely find it helpful to understand the conditions under which it is preferable to add a second treatment rather than switch medications. At least one previous report suggested that clinicians tend to switch medications when there has been minimal response to the treatment and

augment when there has been partial response but whether this approach results in optimal outcomes is not known.[174]

The use of atypical antipsychotics have recently gained prominence in the clinical community. Olanzapine was one of the first atypical antipsychotic medications evaluated. Other atypical antipsychotic medications have since been studied as potential add on, and even monotherapy, treatments for MDD. Aripiprazole has been studied as an add-on therapy for patients not responding to antidepressant medications[77,78,80] but the results are not reported in these figures, as SSRI specific subgroup data is reported only as mean change scores rather than remission and response rates. In the United States, this agent now has an indication as an add-on therapy for patients who do not have adequate response to antidepressant treatment. Similarly, quetiapine , another atypical antipsychotic medication, has been studied as add-on therapy in MDD but the studies were not restricted to SSRI treated patients and the data could not be disaggregated in order to examine the effectiveness of this approach for patients treated only with SSRIs. [175,176]

A recent meta-analysis examined the role for atypical antipsychotic medications as add-on therapy in MDD.[177] The authors reported that the mean odds ratios were similar for the various atypical agents studied (olanzapine, quetiapine, risperidone and aripiprazole); they further reported that they could not appreciate that trial duration or method of establishing treatment resistance influenced the pattern or magnitude of the reported results. The odds ratio reported for response in that meta-analysis (RR = 1.69) is comparable to the RR in the risperidone trial by Mahmoud and colleagues, and more modest than the preliminary results presented by Shelton and colleagues for olanzapine in combination with fluoxetine.[109]

Although lithium was once described as the single agent with the most extensive evidence base for use as an augmenting agent in MDD [178], surveys suggested that it did not have wide uptake in the United States as an agent for treating people with unipolar depression.[179] The results of lithium trials in this evidence review do not support its position as a leading augmentation strategy. We recognize that the trials examined here represent only the portion of lithium trials in which patients were treated with SSRIs initially (meeting the criteria for inclusion in this review). Lithium may, however, have anti-suicide properties[180] or other features that may make it attractive as an add-on agent in some patients with MDD, such as its low potential to induce a mood switch or cycling in depressed patients with strong genetic vulnerability to bipolar disorder.

There is an extremely limited evidentiary base on which to make conclusions regarding the relative efficacies of various combination treatment approaches for patient with an inadequate response to an SSRI. One treatment strategy is to use a combination of treatment modalities, such as a medication in combination with CBT. The STAR*D trial attempted to measure the value of both CBT as monotherapy and CBT in combination with ongoing citalopram treatment, but the number of patients electing CBT or agreeing to the possibility of being randomized to CBT was small compared to the overall sample size, and limits conclusions that can be drawn about CBT from the STAR*D trial. A number of issues related to the provision of CBT in the STAR*D trial have been suggested to account for the relatively small number of patients who found CBT an acceptable option, and these may have limited the generalizability of the patients willing to enter that arm. Another recent trial of a modified cognitive therapy administered to

patients as an augmenting agent following non-response to antidepressant medication found that the psychotherapy was not superior to next step pharmacotherapies that were described to "closely paralleled those in the STAR*D study". A relevant question for clinicians is whether patients who do not have an adequate response to treatment with and antidepressant would do better with an additional medical or with a talk therapy; studies to date do not provide evidence that there are reliable differences in the expected outcomes between these approaches. A caveat to that statement, however, is that patients in both STAR*D and the recent REVAMP trial had many past episodes of depression and it is therefore unknown whether younger patients or those earlier in their course of illness would be more likely to benefit from the addition of CBT than the more chronic group.[127] The TORDIA trial of adolescents with depression who received CBT in combination with medication suggest that this combination might be beneficial in those with a low past illness burden.[17]

Adolescents

Only two trials were identified that were of relevance. The subset of relevant patients could not be extracted from one study[181] (TADS) leaving only one study of children and adolescents that addressed the question of next line treatment for young people who have not had an adequate response to an SSRI. TORDIA appears to emphasize the role for CBT in treating youth. Results continue to emerge from the TORDIA trial[182,183] and will likely provide further information describing the effective components of care for adolescents who have treatment resistant depression.

KQ2. What are the harms of each of the monotherapy or combined therapies among these adults and adolescents? How do the harms compare across different interventions?

It is difficult to summarize any specific trends observed for harms in adults across all the different interventions. In general, the types of the events reported were consistent with the use of antidepressants and these were generally classified as mild with regards to severity. This review has found that the current therapies are likely of modest or uncertain benefit, and in this context evaluation of the harms takes on greater importance. That is, that the margin between the benefits and adverse events associated with the treatment is narrower. Given that many of the treatments that were evaluated likely have equivalent efficacy, evaluation of the harms profile also takes on a greater importance. However, the limitations observed in collecting and reporting harms across the studies, also limits the judgments made regarding the margin of difference between harms and benefits and the relative differences in safety profiles of likely equivalent treatments.

All but one study[91] reported some aspect of safety and tolerability, and similarly, only one study was partially designed to evaluate harms as a primary outcome.[123] No observational studies evaluating this population were eligible for this review and potentially long term consequences of such interventions are therefore not known. Recognizing that from a statistical perspective, it may be difficult to evaluate statistical differences when event rates are low, we found that the

small minority of studies did not undertake such tests when comparing differences between groups. Thus establishing differences in harms profiles rested mainly on judgements.

Rates of discontinuation due to adverse events were variable. In studies with open label prospective failure components, the number of patients who had adverse events who did not proceed to the next phase were not consistently reported. In studies with historical failure, the proportion of subjects who had experienced inadequacy due to intolerance because of harms were not sufficiently detailed. Some studies excluded subjects with any history of drug reactions. Thus, intolerance is not distinguished from inadequate response. Disentangling this issue may prove to be helpful in understanding who may respond in second line treatment.

Washout periods were almost never included in study protocols, and for interventions with switch to new interventions, this may be problematic as very possibly early side effects from these new treatments may reflect symptoms related to withdrawal from the previous SSRI or medication they were taking.

The method of assessing adverse effects differed greatly among studies, with a limited number utilizing standardized scales or methods to assess harms. A priori definitions of serious or severe harms were consistently not specified. Nor was the person who collected reports of harms or their training identified in the majority of studies. Future clinical trials should conform to CONSORT reporting standards for harms.[184]

The single study evaluating harms in adolescents was based on a single well-conducted trial. The harms in this study were collected and reported using relatively unbiased methods.

KQ3. How do these therapies compare in different populations (e.g., different depressive diagnoses, disease severity, ages, gender, racial and socioeconomic group, and medical or psychiatric co-morbidities)? These subgroups will be considered with respect to the different interventions.

There are seven studies in adults with MDD that undertook stratified or subgroup analyses evaluating factors that may impact treatment outcomes in adults and one for adolescents. The findings for baseline severity are inconsistent with two studies suggested no impact and three showing that higher baseline scores were linked to greater change scores or that those with more severe depression are less likely to achieve remission. It is important to note that many studies excluded subjects with psychiatric comorbidities, particularly subjects with anxiety disorders, bipolar disorder and depression with psychotic features. The STAR*D cohort showed that those with concurrent anxiety related psychiatric co-morbidities were less likely to achieve remission and the various treatments did not benefit the anxious group any differently. No clear trend emerges for previous history of failure, age, gender, or race. From a clinical perspective all these factors have face validity as potentially important treatment modifiers. One could argue that there is a greater need to evaluate these characteristics as potential prognostic factors in

populations who have failed to respond could be made. The link with risk factors for predicting failed initial response (first episode) may provide important information for subsequent management of this patient populations.

An analysis of the TORDIA trial[17,95,96] for predictors of treatment response showed that milder depression, less family conflict, and absence of suicidal behavior were associated with greater likelihood of a positive treatment response at 12 weeks.

KQ4. What is the range of recommended clinical actions following the failure of one adequate course of SSRI based on current (<5 years) clinical practice guidelines?

This CER reviewed 23 clinical practice guidelines (CPGs) in the English language and limited to those applicable in the national context or from large national professional associations. There were seven CPGs that were specific only to adolescents, 14 CPG for adults alone, and two applicable to both adults and adolescents. All guidelines were applicable to patients with MDD. From these, four did not provide any recommendations for subjects who had inadequate response. In addition, the three CPGs that included patients with Dysthymia and Subsyndromal depression, did not provide recommendations for those who had failed to respond to previous treatment. To our knowledge a comparison of CPGs on managing MDD, dysthymia and subsyndromal depression does not exist. Moreover, the focus and comparison on evaluation of recommendations for patients who have failed to respond to SSRI is distinct.

From 16 CPGs, 11 did not report any specific definitions for defining adequate response or remission within the guideline. For those CPGs that did report a formal definition of inadequate response, only two CPGs provided clear indications for differential treatment strategies for those with partial versus non-response. This would suggest that the clinical and research community may require both consensus work and knowledge translation strategies to establish acceptable definitions for this group of patients.

Evaluation with the AGREE II instrument showed a consistent difficulty (lowest scores of all domains) in the clarity and consistency of the recommendations; this was amplified for the recommendations targeted at groups that had inadequate responses to initial treatment. This criterion assessed whether the recommendations were clear and unambiguous, such that options were clearly presented, and key recommendations easily identifiable. Most CPG failed to note or include any patient representation (a key stakeholder) in the development process. Although, the CPGs generally rated as acceptable (higher scores) for attempting to link the evidence with the recommendations, the clinical sensibleness of these treatment strategies are not addressed by the AGREE II. A variety of grading systems were used and comparison across different CPGs was problematic. The lack of clear guidance for some treatment options further compounded interpretations across guidelines.

The recommendations for most CPG were stated broadly (switch or augment) and the link between the presentation of the evidence and the specific treatment recommendations was problematic in most CPGs. Few provided clear specification that there was insufficient evidence;

rather any available evidence was summarized with valuations of the strength of the evidence. For adults, the majority of CPGs do not indicate specific types of antidepressants when recommending switching to monotherapy strategies. Increases in the dose and duration of treatment are frequently recommended but the treatment interval or change in dose was not specified. When combined therapy was recommended there was a greater tendency to specify the medication to be added. However, there was great variability in the augmenting agents recommended. The lack of specificity and the relatively high degree of variability is most likely related to the limitations of the evidentiary base.

Guidelines for adolescents scored equally poorly on the AGREE II domain for clarity of presentation of the recommendations. Three from seven guidelines cited adult evidence to justify recommendations for some pharmacological strategies, particularly the use of augmentation agents. Only two CPGs provided definitions of inadequate response. Two guidelines considered failure following non-pharmacological interventions rather than inadequate response to antidepressants. This may reflect a preferential tendency to adopt non-pharmacological strategies in youth. In general, the CPG for adolescents had a greater tendency to specify the medications to consider for both monotherapy and combined therapy. However, all noted the limitations of the evidence applicable to adolescents.

Applicability

The study populations in the eligible literature were relatively homogenous but were limited to predominately white women within a relatively narrow age range, often with mild to moderate depression. Participants in the studies often had many past episodes and high numbers of past treatment failures; it is possible that this represents a selection bias within the studies such that these groups may not reflect the range of patients treated in typical primary care settings. As noted previously, there were few studies evaluating patients with dysthymia and subsyndromal depression and both showed no effects.

The dose range of many of the SSRIs and other treatments were within standard ranges based on product monographs. However, there was some variability in dosing for some augmentation agents and generally a lack of rationale for the selected doses within the study. There is limited data confirming that the doses selected for augmentation agents do, in fact, reflect that optimum doses of those medications in the context of augmenting another agent in a person with MDD.

The variation in treatment duration across studies is potentially problematic; this may reflect lack of consensus as to what constitutes an adequate treatment trial. The outcomes used in most studies (for example, the MADRS or HAM-D) are clinician administered; fewer studies used self report instruments (for example, the QID-SR16 or the PHQ-9). The feasibility of using such instruments in primary care is a consideration that is recognized in the clinical literature. Most of these studies were undertaken in outpatient mental health or primary care settings and are therefore applicable to these contexts.

The TORDIA trial had employed rigorous method monitoring harms (for example, weekly monitoring for those displaying any adverse events) and this may be difficult in replicate primary

care settings. The type of CBT was intensive and had high fidelity, however, it is not clear if accessing therapists with expertise in working with adolescents is feasible in all jurisdictions.

Comparative Effectiveness Review Limitations

This Comparative Effectiveness Review (CER) has several constraints in its methodology in the context of the literature search. Although over 40,000 citations were screened, the citations were limited to those published in the English language. In addition, the search was limited to publications from 1980 forward, as SSRI's were not in use prior to this time. In consultation with the TEP and partners, issues around predictors of response were considered and it was recognized that the scope of the review was sufficiently large to prohibit evaluation of predictors of response.

We identified a large number of studies that had patients who had failed to respond to a variety of antidepressants; those studies that clearly included only 100 percent non-SSRI failed patients, or failed on combination therapies prior to entry into the study were excluded. However, there was a subset of studies that had some proportion who may have failed to SSRI prior to entry. Attempts were made to contact all authors (N = 150 studies) of these studies and asked to provide subgroup data specific to the SSRI failed group. Some authors declared that they could not provide us with stratified analyses and these studies were excluded. The contact information for some authors was incorrect, and several attempts to find information related to the publication investigators was made; for some of these direct contact with the authors was not successful and these studies were excluded as well. Some of these authors did not respond to emails, and after two attempts we excluded these publications as well. A limited number of authors provided us with some stratified results for outcomes and we acknowledge that some of the findings from these stratified analyses may be compromised as the study designs were not such to ensure balance between the SSRI and non-SSRI failed subjects.

A search of the grey literature identified approximately 350 links to regulatory agency documentation, and 171 of these were directly related to any drugs found within our eligible studies. The aim in searching these sources was to identify unpublished trials and for potential deviations for reporting of study findings. However, none of these sources provided additional information to identify unpublished trials and evaluate the potential for reporting biases; this was primarily limited by the population (previous failure to SSRI). Previous research specific to antidepressants has shown significant differences in the information reported to the FDA, relative to the same study publication in peer review journals.[185] Our search of clinical trial registries identified that only 10 from 37 of our eligible studies had been registered.[17,49,76-78,80,95,96,108,113,115,116,119,126,128] Moreover, the abbreviated information within the registry trials was not helpful in identifying selective reporting of outcomes or deviations to the stated protocols. Trial registries are dependent on the investigators to voluntarily update information.

Summary/Conclusions

KQ1.

1) Studies in adults with MDD and an inadequate response to an SSRI included a preponderance of subjects with multiple past depressive episodes and multiple past unsuccessful treatment trials. The generalizability of these data to people with few past episodes of depression may be limited.

2) Studies in adults with MDD and an inadequate response to an SSRI included a high proportion of Whites and women and tended to have an average patient age in the early forties. Studies with more heterogeneous populations with a broader age span are needed. Studies with sufficient sample size to explore whether there are differences in race, sex or across the age spectrum are needed.

3) Studies in adults examining treatment for a major depressive episode following inadequate response to SSRI examined monotherapy compared to monotherapy, monotherapy with combination therapy or combinations of therapies. The majority of studies compared monotherapy (usually ongoing treatment with the initial SSRI) to a combination of therapies (usually ongoing treatment with the initial SSRI in addition to a non-antidepressant medication).
    a. The number of studies comparing single medications against each other (monotherapy compared to monotherapy) following an inadequate response to SSRI is extremely limited. Extant studies are limited in type of agents utilized, sample sizes, and population characteristics. The extant data do not support a difference between various single agent therapies for the outcomes of response and remission.
        i. Strength of Evidence: There is insufficient evidence to evaluate the benefits or harms of switching to a different SSRI, a non-SSRI antidepressant, a non-antidepressant medication or a non-pharmacological treatment following inadequate response to an SSRI.
    b. The largest number of eligible studies examined monotherapy with combination therapy. The majority of studies compared outcomes following ongoing treatment with placebo added to the initial SSRI (the agent to which the subject had not responded by a defined time) to outcomes when an active agent was added to the initial SSRI.
        i. Strength of Evidence: There is insufficient evidence to evaluate the benefits of ongoing monotherapy with an SSRI to combination treatment involving the addition of another antidepressant medication to the initial SSRI.
        ii. Strength of evidence: There is low grade evidence that comparable results are achieved following switch to an alternate antidepressant medication (monotherapy with a new antidepressant) when compared to adding a non-antidepressant treatment to the initial SSRI (traditional augmentation approach).
        iii. Strength of the Evidence: There is moderate and low grade evidence that adding an atypical antipsychotic medication to ongoing SSRI treatment is associated with higher response and remission rates compared to adding a

placebo to ongoing SSRI treatment. There is insufficient evidence for the benefit of other augmentation agents.

c. Studies examining combinations of treatment were also extremely limited in number, types of medications, and homogeneity of populations. Extant data do not suggest that any specific combination of active treatments is superior to another specific combination of treatments.
    i. Strength of Evidence: There is insufficient evidence to evaluate the benefits or harms of specific combinations of treatments relative to alternative combinations.

4) Studies examining response and remission rates in children and adolescents to treatment subsequent to an inadequate SSRI response were extremely limited. Of two potential trials, data could only be extracted from one. This trial was of high quality and the results did not show a difference when monotherapy treatments were compared; switch from the inadequate SSRI to another SSRI was associated with comparable outcomes as switch to an SNRI. The trial did, however, report that combination therapy of a medication and CBT was superior to monotherapy with a medication.
    a. Strength of evidence: There is low-grade evidence to support the use of CBT in combination with an antidepressant following inadequate response to an SSRI for adolescents (age 12-18) with MDD.

5) Studies examining response and remission rates in patients specifically selected to have subsyndromal symptoms associated with inadequate response to SSRI were also extremely limited. Only one trial was eligible and that trial had metabolic parameters as the primary outcomes interest.
    a. Strength of evidence: There is insufficient evidence to support the use of specific treatments for patients with subsyndromal symptoms following an inadequate response to SSRI medications.

6) Studies examining patients with dysthymia (but not MDD) and an inadequate response to an SSRI medication were extremely limited. Only one trial was eligible and that trial did not report a difference between treatment with paroxetine 40mg compared with paroxetine 20mg and amisulpride.
    a. Strength of evidence: There is insufficient evidence to support the use of various treatment approaches for patients with dysthymia who have inadequate response to an SSRI.

KQ2.

1) Harms for interventions used in both adults and adolescents with MDD who had failed to respond to SSRI were derived from RCTs that evaluated treatment strategies in this population; no observational studies were eligible. A clear trend for harms was difficult to specify across the differing interventions in adults. Harms were evaluated and reported in a comprehensive and unbiased manner in the single study in adolescents.

2) Reporting and collecting of harms was problematic, particularly for pre-defining harms including serious and severe events and reporting total number of events per group in study with studies with adults. The single study evaluating harms in adolescents provided high quality evidence for harms within this population when receiving pharmacological and psychological treatment.

KQ3.

1) There is a limited number of studies that have evaluated the impact of disease type, disease severity, previous comorbidities, age, gender and race in studies that have evaluated treatment of adults and adolescents who have failed to respond to an SSRI.

2) Only two studies have evaluated psychiatric comorbidities, and findings from the STAR*D cohort of level two adult patients would suggest that patients with anxiety related disorders (particularly anxious patients) are less likely to achieve remission.

3) There is high quality evidence from the TORDIA trial suggesting that mild depression, less family conflict, and the absence of suicidal behavior is associated with greater likelihood of response in adolescents.

KQ4.

1) The majority of CPGs in adults were applicable to patients with MDD for outpatient and primary care settings. The majority of clinical practice guidelines provided recommendations for patients who had failed previous treatments, but did not specify definitions of "inadequate response".

2) No recommendations for persons with Dysthymia or Subsyndromal depression who had failed previous treatment were found in the limited number of CPGs that included this population.

3) Recommendations for monotherapy, including dose or interval changes, switching to a different SSRI, or a non-SSRI were non-specific as to the drug, interval or dose change.

4) Recommendations for combination therapy tended to recommend specific types of antidepressants and augmenting agents. However, there was inconsistency across CPGs with regards to the types of augmenting agents to use.

Future Research Recommendations

KQ1.

1) Future research should include a broader representation of patients with respect to age and ethnicity. Although more women have a higher prevalence of MDD, future studies should attempt to recruit sufficient men to explore the impact of gender on outcomes of importance. A greater representation of ages, including patients who are younger and older, would provide information that would be of benefit to clinicians.

2) Studies should be more consistent in reporting the manner for determining previous history of failed treatment trials and past episodes of depression. Clinicians could benefit from trials that systematically evaluate response and remission as a function of past number of episodes as treatment responsiveness may vary over the course of illness. Investigators attempting to synthesize the results of multiple studies would also benefit from this information as it is difficult to compare studies with inconsistent results unless this important element of patients' clinical illness is known and considered.

3) Future studies should attempt to determine treatment failure in a prospective manner.

4) A large number of studies included a portion of patients treated with SSRIs and a portion treated with other antidepressant medications. Response and remission rates were not reported as a function of baseline therapy for most studies and although there was a systematic process for contacting authors of the studies, very few were able to provide response and remission data for the sub-group of patients treated with SSRIs. Even registration trials of new add-on treatments for patients not responding to an antidepressant medication have not examined whether the add-on agent is equally effective when added to a range of antidepressant classes. This resulted in the exclusion of a large number of studies containing relevant data because the data for SSRI treated patients could not be disaggregated from those treated with other antidepressants. There appears to be an assumption amongst investigators in this field that response and remission will be comparable regardless of the class of background medication. We are not aware of clinical or neurobiological data to convincingly support this assumption and we suggest that perhaps it could be re-visited. It is likely that the major disincentive to examining this issue is that it will add to required sample sizes if there is a requirement that investigators examine the effectiveness of various add-on therapies as a function of the antidepressant class used as co-therapy. It is possible that if extant studies were examined by disaggregating the various antidepressants employed as the primary treatment, that a preliminary investigation of whether add-on treatments are equally effective for all antidepressant classes could be conducted.

5) Although there are some advantages to employing equipoise designs, there is a need to maintain randomization in studies evaluating interventions in this population.

6) All standard approaches to treating patients with MDD following inadequate response to an SSRI suffer from a lack of adequate evidence to support clinicians' decisions. Whether the

clinician is attempting to optimize the antidepressant medication by changing the dose or duration of the SSRI therapy, switching to a new medication, adding another antidepressant treatment or adding a non-antidepressant agent as augmentation, there is a lack of evidence to guide clinicians and patients in choosing the most appropriate strategy.

7) While the TORDIA trial represents a major advance in the recognition of the need to have data on second line treatment approaches for adolescents with MDD, much more work is required to determine the most effective ways to optimize short and long-term outcomes for adolescents with depression.

8) The apparent benefit of CBT in combination with medication that was observed in the TORDIA trial was not similarly apparent in either the STAR*D or REVAMP trials. The TORDIA trial explicitly states that almost 3 in 4 patients were receiving their first treatment for MDD, while the STAR*D participants had a much higher average burden of illness and the REVAMP participants were specifically chosen to have chronic depression. This raises the question of whether the effectiveness of CBT is determined in part by the illness history and burden of participants; it has been suggested that this might be the case for CBT in patients with bipolar disorder [186]and is worthy of further investigation in unipolar depression. Access to CBT is limited in many jurisdictions and as such, clinicians may choose to reserve the therapy for those who are in a stage of illness where there is a reasonable probability that it will be associated with superior results than other more accessible treatment modalities.

KQ2.

1) Future clinical trials should conform to CONSORT[184] reporting standards for harms. Severe and serious events (including suicidality) were inconsistently reported and improvement in this area is necessary in this area.

KQ3.

1) Future research should endeavor to establish baseline severity in a consistent manner and specify methods and definitions of previous failure.

2) Future research should include a broader representation of patients with respect to age and ethnicity. Although more women have a higher prevalence of MDD, future studies should attempt to recruit sufficient men to explore the impact of gender on outcomes of importance. A greater representation of ages, including patients who are younger and older, would provide information that would be of benefit to clinicians.

KQ4.

1) Development of future CPG for adolescents or adults should provide a clear definition of inadequate response for both pharmacological and non-pharmacological treatments and include standardized methods for establishing this in real world settings.

2) Future CPG development for adolescents and adults should consider adding patient representation in their development process.

3) Future CPG recommendations should provide greater clarity with regards to recommended actions and the link with the evidence. Clinicians should be clear when evidence is insufficient.

Abbreviations

ACP	American College of Physicians
AHRQ	Agency for Healthcare Research and Quality
ASEX	Arizona Sexual Experience Scale
BDI	Beck Depression Inventory
CAM	complementary and alternative medicine
CBT	cognitive behavioral therapy
CCT	controlled clinical trial
CGI	clinical global impressions
CPG	clinical practice guideline
CYP 2D6	enzyme CYP 2D6
DSHS	Department of State Health Services
DSM-IV	Diagnostic and Statistical Manual – 4th edition
ECT	electroconvulsive therapy
GLAD-PC	Guidelines for Adolescent Depression in Primary Care
Ham-D	Hamilton Depression Rating Scale
Ham-D-17/21	17 or 21 questions on the Ham-D scale
HSRProj	Health Services Research Projects in Progress
MADRS	Montgomery Asberg Depression Rating Scale
MDD	major depressive disorder
MDE	major depressive episode
mg/d	milligrams daily
PBO	placebo
PHQ-9	Patient Health Questionnaire 9 item
PICOT	Population, Intervention, Comparator, Outcome, Time-frame
QIDS-SR-16	Quick Inventory of Depressive Symptoms Self Report
QOL	quality of life
RCT	randomized controlled trial
rTMS	repetitive transcranial magnetic stimulation
SD	standard deviation
SNRI	Serotonin–norepinephrine reuptake inhibitors
SOE	strength of evidence
SSRI	selective serotonin reuptake inhibitors
STAR*D	Sequenced Treatment Alternatives to Relieve Depression
TADS	Treatment of Adolescents Study
TEP	technical expert panel
TORDIA	Treatment for Resistant Depression in Adolescents
UKU	Udvalg for Kliniske Undersøgelser (UKU) Side Effect Rating Scale
U.S.	United States
USD	U.S. dollars
USPSTF	United States Preventive Services Task Force
VNS	vagal nerve stimulation
WHO	World Health Organization
WHOQOL	World Health Orgaization Quality of Health Scale

Reference List

1. Canal M, Legangneux E, van Lier JJ, et al. Lack of effect of amisulpride on the pharmacokinetics and safety of lithium. Int J Neuropsychopharmacol 2003;6(2):103-9.
2. World Health Organization. The World Health Report 2001: Mental health: new understanding, new hope. 2001. http://www.who.int/whr/2001/en/
3. Ciechanowski PS, Katon WJ, Russo JE. Depression and diabetes: impact of depressive symptoms on adherence, function, and costs. Arch Intern Med 2000;160(21):3278-85.
4. Aro AR, de Koning HJ, Absetz P, et al. Psychosocial predictors of first attendance for organised mammography screening. J Med Screen 1999;6(2):82-8.
5. McIntyre RS, Soczynska JK, Konarski JZ, et al. The effect of antidepressants on glucose homeostasis and insulin sensitivity: synthesis and mechanisms. Expert Opin Drug Saf 2006;5(1):157-68.
6. Murphy JM, Horton NJ, Monson RR, et al. Cigarette smoking in relation to depression: historical trends from the Stirling County Study. Am J Psychiatry 2003;160(9):1663-9.
7. Van Gool CH, Kempen GI, Penninx BW, et al. Relationship between changes in depressive symptoms and unhealthy lifestyles in late middle aged and older persons: results from the Longitudinal Aging Study Amsterdam. Age Ageing 2003;32(1):81-7.
8. Corcos M, Guilbaud O, Hjalmarsson L, et al. Cytokines and depression: an analogic approach. Biomed Pharmacother 2002;56(2):105-10.
9. Kop WJ, Gottdiener JS, Tangen CM, et al. Inflammation and coagulation factors in persons > 65 years of age with symptoms of depression but without evidence of myocardial ischemia. Am J Cardiol 2002;89(4):419-24.
10. Musselman DL, Miller AH, Porter MR, et al. Higher than normal plasma interleukin-6 concentrations in cancer patients with depression: preliminary findings. Am J Psychiatry 2001;158(8):1252-7.
11. Penninx BW, Kritchevsky SB, Yaffe K, et al. Inflammatory markers and depressed mood in older persons: results from the Health, Aging and Body Composition study. Biol Psychiatry 2003;54(5):566-72.
12. Kop WJ. The integration of cardiovascular behavioral medicine and psychoneuroimmunology: new developments based on converging research fields. Brain Behav Immun 2003;17(4):233-7.
13. Taylor CB, Youngblood ME, Catellier D, et al. Effects of antidepressant medication on morbidity and mortality in depressed patients after myocardial infarction. Arch Gen Psychiatry 2005;62(7):792-8.
14. Wassertheil-Smoller S, Shumaker S, Ockene J, et al. Depression and cardiovascular sequelae in postmenopausal women. The Women's Health Initiative (WHI). Arch Intern Med 2004;164(3):289-98.
15. Gilmour H. Depression and risk of heart disease. Health Reports Vol.19, no.3, 82-003-XPE. Statistics Canada; 2008. http://dsp-psd.pwgsc.gc.ca/collection_2008/statcan/82-003-X/82-003-XIE2008003.pdf
16. Patten SB, Williams JV, Lavorato DH, et al. Major depression as a risk factor for chronic disease incidence: longitudinal analyses in a general population cohort. Gen Hosp Psychiatry 2008;30(5):407-13.
17. Brent D, Emslie G, Clarke G, et al. Switching to another SSRI or to venlafaxine with or without cognitive behavioral therapy for adolescents with SSRI-resistant depression: the TORDIA randomized controlled trial. JAMA 2008;299(8):901-13.
18. Screening and treatment for major depressive disorder in children and adolescents: US Preventive Services Task Force recommendation statement. Pediatrics 2009;123(4):1223-8.
19. Carvalho AF, Cavalcante JL, Castelo MS, et al. Augmentation strategies for treatment-resistant depression: a literature review. J Clin Pharm Ther 2007;32(5):415-28.
20. Souery D, Papakostas GI, Trivedi MH. Treatment-resistant depression. J Clin Psychiatry 2006;67(SUPPL. 6):16-22.
21. McIntyre RS, Fallu A, Konarski JZ. Measurable outcomes in psychiatric disorders: remission as a marker of wellness. Clin Ther 2006;28(11):1882-91.
22. Paykel ES, Ramana R, Cooper Z, et al. Residual symptoms after partial remission: an important outcome in depression. Psychol Med 1995;25(6):1171-80.
23. Higgins JPT. Cochrane Handbook for Systematic Reviews of Interventions Version 5.0.2. 2008. http://www.cochrane-handbook.org/
24. Santaguida P and Raina P. Development of a quality assessment scale specific to harms in studies evaluating the efficacy of health technologies: Manual for using the McHarm. 2010. http://hiru.mcmaster.ca/epc/mcharm.pdf
25. Chou R, Aronson N, Atkins D, et al. AHRQ series paper 4: assessing harms when comparing medical interventions: AHRQ and the effective health-care program. J Clin Epidemiol 2010;63(5):502-12.

26. The AGREE Next Steps Consortium. Appraisal of guidelines for research and evaluation II (AGREE II). The AGREE Research Trust; 2009. http://fhswedge.csu.mcmaster.ca/pebc/agreetrust/AGREE%20II%20-%20User's%20Manual%20and%2023-item%20Instrument-%20ENGLISH.pdf
27. GRADE working group. Grading the quality of evidence and the strength of recommendations. www gradeworkinggroup org 2009. 2009.
28. Owens DK, Lohr KN, Atkins D, et al. AHRQ series paper 5: grading the strength of a body of evidence when comparing medical interventions--agency for healthcare research and quality and the effective health-care program. J Clin Epidemiol 2010;63(5):513-23.
29. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders IV. 1994. http://allpsych.com/disorders/dsm.html
30. Kessler RC, Berglund P, Demler O, et al. The epidemiology of major depressive disorder - Results from the National Comorbidity Survey Replication (NCS-R). Journal of the American Medical Association 2003;289(23):3095-105.
31. Alonso J, Angermeyer MC, Bernert S, et al. Prevalence of mental disorders in Europe: results from the European Study of the Epidemiology of Mental Disorders (ESEMeD) project. Acta Psychiatr Scand Suppl 2004;(420):21-7.
32. Beck CA, Patten SB. Adjustment to antidepressant utilization rates to account for depression in remission. Compr Psychiatry 2004;45(4):268-74.
33. Brugha T, Jenkins R, Bebbington P, et al. Risk factors and the prevalence of neurosis and psychosis in ethnic groups in Great Britain. Soc Psychiatry Psychiatr Epidemiol 2004;39(12):939-46.
34. Kessler RC, Demler O, Frank RG, et al. Prevalence and treatment of mental disorders, 1990 to 2003. N Engl J Med 2005;352(24):2515-23.
35. Gonzalez HM, Vega WA, Williams DR, et al. Depression care in the United States: too little for too few. Arch Gen Psychiatry 2010;67(1):37-46.
36. Spijker J, de Graaf R, Bijl RV, et al. Duration of major depressive episodes in the general population: results from The Netherlands Mental Health Survey and Incidence Study (NEMESIS). Br J Psychiatry 2002;181:208-13.
37. Birmaher B, Brent D, Bernet W, et al. Practice parameter for the assessment and treatment of children and adolescents with depressive disorders. J Am Acad Child Adolesc Psychiatry 2007;46(11):1503-26.
38. Birmaher B, Ryan ND, Williamson DE, et al. Childhood and adolescent depression: A review of the past 10 years, Part II. J Am Acad Child Adolesc Psychiatry 1996;35(12):1575-83.
39. Murray CJ, Lopez AD. Evidence-based health policy--lessons from the Global Burden of Disease Study. Science 1996;274(5288):740-3.
40. The Standing Senate Committee on Social Affairs SaT. Out of the shadows at last: transforming mental health, mental illness and addiction services in Canada. 2006. http://www.parl.gc.ca/39/1/parlbus/commbus/senate/com-e/soci-e/rep-e/rep02may06-e.htm
41. Rapaport MH, Clary C, Fayyad R, et al. Quality-of-life impairment in depressive and anxiety disorders. Am J Psychiatry 2005;162(6):1171-8.
42. Tompson MC, Pierre CB, Boger KD, et al. Maternal depression, maternal expressed emotion, and youth psychopathology. J Abnorm Child Psychol 2010;38(1):105-17.
43. Eckshtain D, Ellis DA, Kolmodin K, et al. The effects of parental depression and parenting practices on depressive symptoms and metabolic control in urban youth with insulin dependent diabetes. J Pediatr Psychol 2010;35(4):426-35.
44. Ramchandani P, Stein A. The impact of parental psychiatric disorder on children: avoiding stigma, improving care [letter]. BMJ: British Medical Journal 2003;327(7409):242-3.
45. Adler DA, McLaughlin TJ, Rogers WH, et al. Job performance deficits due to depression. Am J Psychiatry 2006;163(9):1569-76.
46. Lerner D, Adler DA, Chang H, et al. The clinical and occupational correlates of work productivity loss among employed patients with depression. J Occup Environ Med 2004;46(6 Suppl):S46-S55
47. Hoge CW, Lesikar SE, Guevara R, et al. Mental disorders among U.S. military personnel in the 1990s: association with high levels of health care utilization and early military attrition. Am J Psychiatry 2002;159(9):1576-83.
48. Williams JW, Jr., Pignone M, Ramirez G, et al. Identifying depression in primary care: a literature synthesis of case-finding instruments. Gen Hosp Psychiatry 2002;24(4):225-37.
49. Rush AJ, Trivedi MH, Wisniewski SR, et al. Bupropion-SR, sertraline, or venlafaxine-XR after failure of SSRIs for depression. N Engl J Med 2006;354(12):1231-42.
50. Bilsker D, Goldner EM, Jones W. Health service patterns indicate potential benefit of supported self-management for depression in primary care. Can J Psychiatry 2007;52(2):86-95.
51. Kupfer DJ. Long-term treatment of depression. Journal of Clinical Psychiatry 1991;52(5 SUPPL.):28-34.
52. Papakostas GI, Perlis RH, Scalia MJ, et al. A meta-analysis of early sustained response rates

between antidepressants and placebo for the treatment of major depressive disorder. J Clin Psychopharmacol 2006;26(1):56-60.
53. Posternak MA, Zimmerman M. Is there a delay in the antidepressant effect? A meta-analysis. J Clin Psychiatry 2005;66(2):148-58.
54. Taylor MJ, Freemantle N, Geddes JR, et al. Early onset of selective serotonin reuptake inhibitor antidepressant action: systematic review and meta-analysis. Arch Gen Psychiatry 2006;63(11):1217-23.
55. Wade A, Friis AH. The onset of effect for escitalopram and its relevance for the clinical management of depression. Curr Med Res Opin 2006;22(11):2101-10.
56. Trivedi MH, Greer TL, Grannemann BD, et al. Exercise as an augmentation strategy for treatment of major depression. J Psychiatr Pract 2006;12(4):205-13.
57. Parikh SV, Segal ZV, Grigoriadis S, et al. Canadian Network for Mood and Anxiety Treatments (CANMAT) Clinical guidelines for the management of major depressive disorder in adults. II. Psychotherapy alone or in combination with antidepressant medication. J Affect Disord 2009;117(SUPPL. 1):S15-S25
58. Ravindran AV, Lam RW, Filteau MJ, et al. Canadian Network for Mood and Anxiety Treatments (CANMAT) Clinical guidelines for the management of major depressive disorder in adults. V. Complementary and alternative medicine treatments. J Affect Disord 2009;117(SUPPL. 1):S54-S64
59. Qaseem A, Snow V, Denberg TD, et al. Using second-generation antidepressants to treat depressive disorders: a clinical practice guideline from the American College of Physicians. Ann Intern Med 2008;149(10):725-33.
60. Hamilton M. A rating scale for depression. J Neurol Neurosurg Psychiatry 1960;23:56-62.
61. Montgomery SA, Asberg M. A new depression scale designed to be sensitive to change. Br J Psychiatry 1979;134:382-9.
62. Uher R, Farmer A, Maier W, et al. Measuring depression: comparison and integration of three scales in the GENDEP study. Psychol Med 2008;38(2):289-300.
63. Riedel M, Moller HJ, Obermeier M, et al. Response and remission criteria in major depression - A validation of current practice. J Psychiatr Res 2010;
64. Gartlehner, G, Hansen, R. A., Thieda, P., et al. Comparative Effectiveness of Second-Generation Antidepressants in the Pharmacologic Treatment of Adult Depression. Comparative Effectiveness Review Number 7. www.effectivehealthcare.ahrq.gov/reports/final.cfm.
65. Williams AL, Girard C, Jui D, et al. S-adenosylmethionine (SAMe) as treatment for depression: a systematic review. [Review] [37 refs]. Clin Invest Med 2005;28(3):132-9.
66. Souery D, Amsterdam J, de Montigny C, et al. Treatment resistant depression: methodological overview and operational criteria. [Review] [52 refs]. Eur Neuropsychopharmacol 1999;9(1-2):83-91.
67. Berlim MT, Fleck MP, Turecki G. Current trends in the assessment and somatic treatment of resistant/refractory major depression: An overview. Ann Med 2008;40(2):149-59.
68. Fava M. Diagnosis and definition of treatment-resistant depression. Biological Psychiatry 2003;53(8):649-59.
69. Frank E, Prien RF, Jarrett RB, et al. Conceptualization and Rationale for Consensus Definitions of Terms in Major Depressive Disorder - Remission, Recovery, Relapse, and Recurrence. Arch Gen Psychiatry 1991;48(9):851-5.
70. Hartling L, Ospina M, Liang Y, et al. Risk of bias versus quality assessment of randomised controlled trials: cross sectional study. BMJ: British Medical Journal 2009;339:b4012
71. Wells GA, Shea B, O'Connell D, et al. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. Last Accessed February 1, 2009. http://www.ohri.ca/programs/clinical_epidemiology/oxford.htm.
72. Heres S, Davis J, Maino K, et al. Why olanzapine beats resperidone, resperidone beats quetiapine, and quetiapine beats olanzapine: An exploratory analysis of head-to-head comparison studies of second-generation antipsychotics. Am J Psychiatry 2006;163(2):185-94.
73. Blumenthal D, Campbell E, Anderson M, et al. Withholding research results in academic life science - Evidence from a national survey of faculty. JAMA 1997;277(15):1224-8.
74. Wiles NJ, Hollinghurst S, Mason V, et al. A randomized controlled trial of cognitive behavioural therapy as an adjunct to pharmacotherapy in primary care based patients with treatment resistant depression: A pilot study. Behav Cogn Psychother 2008;36(1):21-33.
75. Keitner GI, Garlow SJ, Ryan CE, et al. A randomized, placebo-controlled trial of risperidone augmentation for patients with difficult-to-treat unipolar, non-psychotic major depression. [References]. J Psychiatr Res 2009;43(3):205-14.
76. Mahmoud RA, Pandina GJ, Turkoz I, et al. Risperidone for treatment-refractory major depressive disorder: a randomized trial.[summary for patients in Ann Intern Med. 2007 Nov 6;147(9):I34; PMID: 17975178]. Ann Intern Med 2007;147(9):593-602.
77. Berman RM, Marcus RN, Swanink R, et al. The efficacy and safety of aripiprazole as

adjunctive therapy in major depressive disorder: a multicenter, randomized, double-blind, placebo-controlled study.. J Clin Psychiatry 2007;68(6):843-53.
78. Marcus RN, McQuade RD, Carson WH, et al. The efficacy and safety of aripiprazole as adjunctive therapy in major depressive disorder: a second multicenter, randomized, double-blind, placebo-controlled study. J Clin Psychopharmacol 2008;28(2):156-65.
79. Berman RM, Fava M, Thase ME, et al. Aripiprazole augmentation in major depressive disorder: a double-blind, placebo-controlled study in patients with inadequate response to antidepressants. Cns Spectrums 2009;14(4):197-206.
80. Thase ME, Trivedi MH, Nelson JC, et al. Examining the efficacy of adjunctive aripiprazole in major depressive disorder: A pooled analysis of 2 studies. Primary Care Companion to the Journal of Clinical Psychiatry 2008;10(6):440-7.
81. Carta MG, Hardoy MC, Pilu A, et al. Improving physical quality of life with group physical activity in the adjunctive treatment of major depressive disorder. Clinical Practice and Epidemiology in Mental Health 2008;4, 2008. Article Number: 1. Date of Publication: 26 Jan 2008.:
82. McGrath PJ, Stewart JW, Fava M, et al. Tranylcypromine versus venlafaxine plus mirtazapine following three failed antidepressant medication trials for depression: a STAR*D report. Am J Psychiatry 2006;163(9):1531-41.
83. Fava M, Rush AJ, Wisniewski SR, et al. A comparison of mirtazapine and nortriptyline following two consecutive failed medication treatments for depressed outpatients: a STAR*D report. Am J Psychiatry 2006;163(7):1161-72.
84. Kenney C, Hunter C, Mejia N, et al. Is history of depression a contraindication to treatment with tetrabenazine? Clin Neuropharmacol 2006;29(5):259-64.
85. Perlis RH, Moorjani P, Fagerness J, et al. Pharmacogenetic analysis of genes implicated in rodent models of antidepressant response: association of TREK1 and treatment resistance in the STAR(*)D study. Neuropsychopharmacology 2008;33(12):2810-9.
86. Yates WR, Mitchell J, Rush AJ, et al. Clinical features of depressed outpatients with and without co-occurring general medical conditions in STAR*D. Gen Hosp Psychiatry 2004;26(6):421-9.
87. Lesser IM, Leuchter AF, Trivedi MH, et al. Characteristics of insured and noninsured outpatients with depression in STAR*D. Psychiatr Serv 2005;56(8):995-1004.
88. Perlis RH, Patrick A, Smoller JW, et al. When is pharmacogenetic testing for antidepressant response ready for the clinic? A cost-effectiveness analysis based on data from the STAR*D study. Neuropsychopharmacology 2009;34(10):2227-36.
89. Corya SA, Williamson D, Sanger TM, et al. A randomized, double-blind comparison of olanzapine/fluoxetine combination, olanzapine, fluoxetine, and venlafaxine in treatment-resistant depression. Depress Anxiety 2006;23(6):364-72.
90. Shelton RC, Williamson DJ, Corya SA, et al. Olanzapine/fluoxetine combination for treatment-resistant depression: a controlled study of SSRI and nortriptyline resistance. J Clin Psychiatry 2005;66(10):1289-97.
91. Bondolfi G, Aubry JM, Golaz J, et al. A stepwise drug treatment algorithm to obtain complete remission in depression: a Geneva study. Swiss Medical Weekly 2006;136(5-6):78-85.
92. Nierenberg AA, Fava M, Trivedi MH, et al. A comparison of lithium and T(3) augmentation following two failed medication treatments for depression: a STAR*D report. Am J Psychiatry 2006;163(9):1519-30.
93. Lin CH, Lin KS, Lin CY, et al. Time to rehospitalization in patients with major depressive disorder taking venlafaxine or fluoxetine. J Clin Psychiatry 2008;69(1):54-9.
94. Rapaport MH, Gharabawi GM, Canuso CM, et al. Effects of risperidone augmentation in patients with treatment-resistant depression: Results of open-label treatment followed by double-blind continuation. Neuropsychopharmacology 2006;31(11):2505-13.
95. Asarnow JR, Emslie G, Clarke G, et al. Treatment of selective serotonin reuptake inhibitor-resistant depression in adolescents: Predictors and moderators of treatment response. [References]. J Am Acad Child Adolesc Psychiatry 2009;48(3):330-9.
96. Brent DA, Emslie GJ, Clarke GN, et al. Predictors of spontaneous and systematically assessed suicidal adverse events in the treatment of SSRI-resistant depression in adolescents (TORDIA) study. [References]. The American journal of psychiatry 2009;166(4):418-26.
97. Vitiello B, Brent DA, Greenhill LL, et al. Depressive Symptoms and Clinical Status During the Treatment of Adolescent Suicide Attempters (TASA) Study. Journal of the American Academy of Child and Adolescent Psychiatry 2009;48(10):997-1004.
98. Brent DA, Greenhill LL, Compton S, et al. The Treatment of Adolescent Suicide Attempters Study (TASA): Predictors of Suicidal Events in an Open Treatment Trial. J

Am Acad Child Adolesc Psychiatry 2009;48(10):987-96.
99. Stanley B, Brown G, Brent DA, et al. Cognitive-Behavioral Therapy for Suicide Prevention (CBT-SP): Treatment Model, Feasibility, and Acceptability. J Am Acad Child Adolesc Psychiatry 2009;48(10):1005-13.
100. Zourkova A. Effect of mirtazapine and paroxetine on residual symptoms of depressive disorders and their effect on P450 CYP 2D6 activity. Homeost Health Dis 2001;41(6):242-9.
101. Rocca P, Marchiaro L, Rasetti R, et al. A comparison of paroxetine versus paroxetine plus amisulpride in the treatment of dysthymic disorder: efficacy and psychosocial outcomes. Psychiatry Res 2002;112(2):145-52.
102. Alexopoulos GS, Canuso CM, Gharabawi GM, et al. Placebo-controlled study of relapse prevention with risperidone augmentation in older patients with resistant depression. Am J Geriatr Psychiatry 2008;16(1):21-30.
103. Rush AJ, Wisniewski SR, Warden D, et al. Selecting among second-step antidepressant medication monotherapies: predictive value of clinical, demographic, or first-step treatment features. Arch Gen Psychiatry 2008;65(8):870-80.
104. Fava M, Rush AJ, Alpert JE, et al. Difference in treatment outcome in outpatients with anxious versus nonanxious depression: A STAR*D report. Am J Psychiatry 2008;165(3):342-51.
105. Warden D, Rush AJ, Wisniewski SR, et al. What predicts attrition in second step medication treatments for depression?: a STAR*D Report. Int J Neuropsychopharmacol 2009;12(4):459-73.
106. Perlis RH, Iosifescu DV, Alpert J, et al. Effect of medical comorbidity on response to fluoxetine augmentation or dose increase in outpatients with treatment-resistant depression. Psychosomatics 2004;45(3):224-9.
107. Perlis RH, Alpert J, Nierenberg AA, et al. Clinical and sociodemographic predictors of response to augmentation, or dose increase among depressed outpatients resistant to fluoxetine 20 mg/day. Acta Psychiatr Scand 2003;108(6):432-8.
108. Thase ME, Corya SA, Osuntokun O, et al. A randomized, double-blind comparison of olanzapine/fluoxetine combination, olanzapine, and fluoxetine in treatment-resistant major depressive disorder. J Clin Psychiatry 2007;68(2):224-36.
109. Shelton RC, Tollefson GD, Tohen M, et al. A novel augmentation strategy for treating resistant major depression. Am J Psychiatry 2001;158(1):131-4.
110. Licht RW, Qvitzau S. Treatment strategies in patients with major depression not responding to first-line sertraline treatment. A randomised study of extended duration of treatment, dose increase or mianserin augmentation. Psychopharmacology (Berl) 2002;161(2):143-51.
111. Ferreri M, Lavergne F, Berlin I, et al. Benefits from mianserin augmentation of fluoxetine in patients with major depression non-responders to fluoxetine alone. Acta Psychiatr Scand 2001;103(1):66-72.
112. Rush AJ, Warden D, Wisniewski SR, et al. STAR*D: revising conventional wisdom. CNS Drugs 2009;23(8):627-47.
113. Thase ME, Friedman ES, Biggs MM, et al. Cognitive therapy versus medication in augmentation and switch strategies as second-step treatments: a STAR*D report. Am J Psychiatry 2007;164(5):739-52.
114. Rush AJ, Trivedi MH, Wisniewski SR, et al. Acute and longer-term outcomes in depressed outpatients requiring one or several treatment steps: a STAR*D report. Am J Psychiatry 2006;163(11):1905-17.
115. Perahia DG, Quail D, Desaiah D, et al. Switching to duloxetine from selective serotonin reuptake inhibitor antidepressants: a multicenter trial comparing 2 switching techniques. J Clin Psychiatry 2008;69(1):95-105.
116. Trivedi MH, Fava M, Wisniewski SR, et al. Medication augmentation after the failure of SSRIs for depression. N Engl J Med 2006;354(12):1243-52.
117. Lynch TR, Cheavens JS, Cukrowicz KC, et al. Treatment of older adults with co-morbid personality disorder and depression: A dialectical behavior therapy approach. Int J Geriatr Psychiatry 2007;22(2):131-43.
118. Sokolski KN, Conney JC, Brown BJ, et al. Once-daily high-dose pindolol for SSRI-refractory depression. Psychiatry Res 2004;125(2):81-6.
119. Michelson D, Adler LA, Amsterdam JD, et al. Addition of atomoxetine for depression incompletely responsive to sertraline: a randomized, double-blind, placebo-controlled study. J Clin Psychiatry 2007;68(4):582-7.
120. Landén M, Björling G, Agren H, et al. A randomized, double-blind, placebo-controlled trial of buspirone in combination with an SSRI in patients with treatment-refractory depression. J Clin Psychiatry 1998;59(12):664-8.
121. Landen M, Hogberg P, Thase ME. Incidence of sexual side effects in refractory depression during treatment with citalopram or paroxetine. J Clin Psychiatry 2005;66(1):100-6.
122. Landén M, Eriksson E, Agren H, et al. Effect of buspirone on sexual dysfunction in depressed patients treated with selective serotonin reuptake inhibitors. J Clin Psychopharmacol 1999;19(3):268-71.

123. Fava M, Thase ME, DeBattista C. A multicenter, placebo-controlled study of modafinil augmentation in partial responders to selective serotonin reuptake inhibitors with persistent fatigue and sleepiness. J Clin Psychiatry 2005;66(1):85-93.
124. Fava M, Alpert J, Nierenberg A, et al. Double-blind study of high-dose fluoxetine versus lithium or desipramine augmentation of fluoxetine in partial responders and nonresponders to fluoxetine. J Clin Psychopharmacol 2002;22(4):379-87.
125. Appelberg BG, Syvalahti EK, Koskinen TE, et al. Patients with severe depression may benefit from buspirone augmentation of selective serotonin reuptake inhibitors: results from a placebo-controlled, randomized, double-blind, placebo wash-in study. J Clin Psychiatry 2001;62(6):448-52.
126. Preskorn SH, Baker B, Kolluri S, et al. An innovative design to establish proof of concept of the antidepressant effects of the NR2B subunit selective N-methyl-D-aspartate antagonist, CP-101,606, in patients with treatment-refractory major depressive disorder. J Clin Psychopharmacol 2008;28(6):631-7.
127. George TP, Sacco KA, Vessicchio JC, et al. Nicotinic antagonist augmentation of selective serotonin reuptake inhibitor-refractory major depressive disorder: a preliminary study. J Clin Psychopharmacol 2008;28(3):340-4.
128. Shapira B, Nemets B, Trachtenberg A, et al. Phenytoin as an augmentation for SSRI failures: a small controlled study. J Affect Disord 2006;96(1-2):123-6.
129. Seidman SN, Miyazaki M, Roose SP. Intramuscular testosterone supplementation to selective serotonin reuptake inhibitor in treatment-resistant depressed men: randomized placebo-controlled clinical trial. J Clin Psychopharmacol 2005;25(6):584-8.
130. Baumann P, Nil R, Souche A, et al. A double-blind, placebo-controlled study of citalopram with and without lithium in the treatment of therapy-resistant depressive patients: a clinical, pharmacokinetic, and pharmacogenetic investigation. J Clin Psychopharmacol 1996;16(4):307-14.
131. Altamura AC, Dell'Osso B, Buoli M, et al. Intravenous augmentative citalopram versus clomipramine in partial/nonresponder depressed patients: a short-term, low dose, randomized, placebo-controlled study. J Clin Psychopharmacol 2008;28(4):406-10.
132. Altamura AC, Dell'Osso B, Buoli M, et al. Short-term intravenous citalopram augmentation in partial/nonresponders with major depression: a randomized placebo-controlled study. International Clinical Psychopharmacology 2008;23(4):198-202.
133. Perry EB, Berman RM, Sanacora G, et al. Pindolol augmentation in depressed patients resistant to selective serotonin reuptake inhibitors: a double-blind, randomized, controlled trial. J Clin Psychiatry 2004;65(2):238-43.
134. Nemets B, Mishory A, Levine J, et al. Inositol addition does not improve depression in SSRI treatment failures. J Neural Transm 1999;106(7-8):795-8.
135. Fava M, Rosenbaum JF, McGrath PJ, et al. Lithium and tricyclic augmentation of fluoxetine treatment for resistant major depression: a double-blind, controlled study. The American journal of psychiatry 1994;151(9):1372-4.
136. Dunner DL, Amsterdam JD, Shelton RC, et al. Efficacy and tolerability of adjunctive ziprasidone in treatment-resistant depression: a randomized, open-label, pilot study. J Clin Psychiatry 2007;68(7):1071-7.
137. Teghtsoonian K. Depression and mental health in neoliberal times: A critical analysis of policy and discourse. Soc Sci Med 2009;69(1):28-35.
138. Omega-3 might ease depression, enhance mind & memory. There's a growing body of evidence that DHA benefits gray matter, but it's not a cure-all, say researchers. Health News 2007;13(6):4-5.
139. Goldberg JF, Gerstein RK, Wenze SJ, et al. Dysfunctional attitudes and cognitive schemas in bipolar manic and unipolar depressed outpatients: implications for cognitively based psychotherapeutics. J NERV MENT DIS 2008;196(3):207-10.
140. Reimherr FW, Strong RE, Marchant BK, et al. Factors affecting return of symptoms 1 year after treatment in a 62-week controlled study of fluoxetine in major depression. J Clin Psychiatry 2001;62:Suppl-23
141. Lafer B, Fava M, Hammerness P, et al. The influence of DST and TRH test administration on depression assessments: a controlled study. Biological Psychiatry 1993;34(9):650-3.
142. Dinan TG. Lithium augmentation in sertraline-resistant depression: a preliminary dose-response study. Acta Psychiatr Scand 1993;88(4):300-1.
143. Nefazodone and liver damage. Harv Ment Health Lett 2002;19(2):8
144. Vieweg WV, Wood MA. Tricyclic antidepressants, QT interval prolongation, and torsade de pointes. Psychosomatics 2004;45(5):371-7.
145. Zuckerbrot R, Cheung M.H, Jensen P, et al. Guidelines for adolescent Depression in primary care (GLAD-PC) I, Identification, assessment, and initial management. Official Journal of the American Academy of Pediatrics 2009;120(5):1299-312.
146. Cheung AH, Zuckerbrot RA, Jensen PS, et al. Guidelines for adolescent depression in primary care (GLAD-PC): II. Treatment and

ongoing management. Pediatrics 2007;120(5):e1313-e1326
147. Dudley M, Hadzi-Pavlovic D, Andrews D, et al. New-generation antidepressants, suicide and depressed adolescents: how should clinicians respond to changing evidence?. [Review] [107 refs]. AUST NZ J PSYCHIATRY 2008;42(6):456-66.
148. Hughes CW, Emslie GJ, Crismon ML, et al. Texas Children's Medication Algorithm Project: update from Texas Consensus Conference Panel on Medication Treatment of Childhood Major Depressive Disorder. [131 refs]. J Am Acad Child Adolesc Psychiatry 2007;46(6):667-86.
149. Gallagher R. Evidence-based psychotherapies for depressed adolescents: A review and clinical guidelines. [References]. Primary Psychiatry 2005;12(9):33-9.
150. National Institute for Clinical Excellence. Depression in Children and Young People : Identification and management in primary, community and secondary care. 2005:1-233. 2005.
151. Jaehne, M. E. Health Care Guideline: Major Depression in Adults in Primary Care 12th Edition. Institute for Clinical Systems Improvement.
152. Horsley L. ACP guideline on second-generation antidepressants for depression treatment. Am Fam Physician 2009;80(3):291-4.
153. Bryan Karasu, Alan Gelenberg, Arnold Merriam, et al. Practice guideline for treatment of patients with depression disorder second ediдtion. APA Practice Guidlines 2009;1-78.
154. Depression clinical practice guidlines. National Guidline Clearinghouse 2004;1-20.
155. Anderson IM, Ferrier IN, Baldwin RC, et al. Evidence-based guidelines for treating depressive disorders with antidepressants: a revision of the 2000 British Association for Psychopharmacology guidelines. [620 refs]. J Psychopharmacol (Oxf) 2008;22(4):343-96.
156. Steinman LE, Frederick JT, Prohaska T, et al. Recommendations for treating depression in community-based older adults. [52 refs]. Am J Prev Med 2007;33(3):175-81.
157. Bauer M, Bschor T, Pfennig A, et al. World Federation of Societies of Biological Psychiatry (WFSBP) Guidelines for Biological Treatment of Unipolar Depressive Disorders in Primary Care. [Review] [380 refs]. World Journal of Biological Psychiatry 2007;8(2):67-104.
158. Davidson KW, Kupfer DJ, Bigger JT, et al. Assessment and treatment of depression in patients with cardiovascular disease: National Heart, Lung, and Blood Institute Working Group Report. [25 refs]. Psychosom Med 2006;68(5):645-50.
159. Ellis P, Royal Australian and New Zealand College of Psychiatrists Clinical Practice Guidelines Team for Depression. Australian and New Zealand clinical practice guidelines for the treatment of depression. [Review] [247 refs]. AUST NZ J PSYCHIATRY 2004;38(6):389-407.
160. Malhi GS, Adams D, Porter R, et al. Clinical practice recommendations for depression. Acta Psychiatr Scand 2009;119(SUPPL. 439):8-26.
161. Lam RW, Kennedy SH, Grigoriadis S, et al. Canadian Network for Mood and Anxiety Treatments (CANMAT) Clinical guidelines for the management of major depressive disorder in adults. III. Pharmacotherapy. J Affect Disord 2009;117(SUPPL. 1):S26-S43
162. Ravindran AV, Lam RW, Filteau MJ, et al. Canadian Network for Mood and Anxiety Treatments (CANMAT) Clinical guidelines for the management of major depressive disorder in adults. V. Complementary and alternative medicine treatments. J Affect Disord 2009;117:S54-S64
163. Kennedy SH, Lam RW, Parikh SV, et al. Canadian Network for Mood and Anxiety Treatments (CANMAT) Clinical guidelines for the management of major depressive disorder in adults Introduction. J Affect Disord 2009;117:S1-S2
164. Anderson IM, Haddad PM. CANMAT Guidelines for depression: Clear and user-friendly. J Affect Disord 2009;117:S3-S4
165. Conn DK, Gibson M, Feldman S, et al. National guidelines for seniors' mental health: The Assessment and Treatment of Mental Health Issues in Long-Term Care Homes (focus on mood and behaviour symptoms). Canadian Journal of Geriatrics 2006;9(SUPPL. 2):S59-S64
166. R Mahendran, H L Yap. Clinical practice guidelines for depression. Singapore Med J 200546(11):610-5. 2005.
167. National Institute for Health and Clinical Excellence. Depression: the treatment and management of depression in adults. 2009:1-585. 2009.
168. National Institute for Health and Clinical Excellence. Computerised cognitive behaviour therapy for depression and anxiety. 200897. 2008.
169. New Zealand Guidelines Group. Identification of Common Mental Disorders and Management of Depression in Primary Care. 2008:1-190. 2008.
170. Ravindran AV. If a patient does not respond to a full dose of fluvoxamine for at least 12 weeks, what alternatives should be considered? J PSYCHIATRY NEUROSCI 1998;23(2):136
171. national clinical practice, guideline. Depression: Management of depression in

primary and secondary care. national clinical practice guideline number 23 2009;23:1-363.
172. Carolusson S. Hypnosis and transference in the treatment of depression. Hypnos 1998;25(2):78-86.
173. Birmaher B, Brent D, Bernet W, et al. Practice parameter for the assessment and treatment of children and adolescents with depressive disorders. J Am Acad Child Adolesc Psychiatry 2007;46(11):1503-26.
174. Papakostas GI, Petersen TJ, Green C, et al. A description of next-step switching versus augmentation practices for outpatients with treatment-resistant major depressive disorder enrolled in an academic specialty clinic. Ann Clin Psychiatry 2005;17(3):161-5.
175. Blier P SST. Potential mechanisms of action of atypical antipsychotic medications in treatment-resistant depression and anxiety. J Clin Psychiatry 2005;66(Supplement 8):30-40.
176. McIntyre RS, Muzina DJ, Adams A, et al. Quetiapine XR efficacy and tolerability as monotherapy and as adjunctive treatment to conventional antidepressants in the acute and maintenance treatment of major depressive disorder: a review of registration trials. Expert Opin Pharmacother 2009;10(18):3061-75.
177. Nelson JC, Papakostas GI. Atypical antipsychotic augmentation in major depressive disorder: a meta-analysis of placebo-controlled randomized trials. Am J Psychiatry 2009;166(9):980-91.
178. Bauer M, Pretorius HW, Constant EL, et al. Extended-release quetiapine as adjunct to an antidepressant in patients with major depressive disorder: results of a randomized, placebo-controlled, double-blind study. J Clin Psychiatry 2009;70(4):540-9.
179. Fredman SJ, Fava M, Kienke AS, et al. Partial response, nonresponse, and relapse with selective serotonin reuptake inhibitors in major depression: a survey of current "next-step" practices. J Clin Psychiatry 2000;61(6):403-8.
180. Baldessarini RJ, Tondo L, Hennen J. Treating the suicidal patient with bipolar disorder. Reducing suicide risk with lithium. Ann N Y Acad Sci 2001;932:24-38.
181. Grudell ABM, Sweetser S, Camilleri M, et al. A Controlled Pharmacogenetic Trial of Sibutramine on Weight Loss and Body Composition in Obese or Overweight Adults. Gastroenterology 2008;135(4):1142-54.
182. Emslie GJ, Mayes T, Porta G, et al. Treatment of Resistant Depression in Adolescents (TORDIA): week 24 outcomes. Am J Psychiatry 2010;167(7):782-91.
183. Kennard BD, Clarke GN, Weersing VR, et al. Effective components of TORDIA cognitive-behavioral therapy for adolescent depression: preliminary findings. Journal of Consulting & Clinical Psychology 2009;77(6):1033-41.
184. Ioannidis JPA, Evans SJW, Gotzsche PC, et al. Better reporting of harms in randomized trials: An extension of the CONSORT statement. Ann Intern Med 2004;141(10):781-8.
185. Turner EH, Matthews AM, Linardatos E, et al. Selective publication of antidepressant trials and its influence on apparent efficacy. N Engl J Med 2008;358(3):252-60.
186. Scott J, Paykel E, Morriss R, et al. Cognitive-behavioural therapy for severe and recurrent bipolar disorders: randomised controlled trial. BR J PSYCHIATRY 2006;188:313-20.